



جامعة غليزان

كلية العلوم الإقتصادية، التجارية وعلوم التسيير
قسم العلوم التجارية

مطبوعة بيداغوجية بعنوان :

الإحصاء 4

موجهة لطلبة السنة الثانية ليسانس علوم التسيير / ميدان العلوم الاقتصادية والتجارية وعلوم التسيير

من إعداد

د. مزواغي جيلالي

السنة الجامعية: 2026/2025

كلية العلوم الاقتصادية، التجارية وعلوم التسيير
المجلس العلمي للكلية
رقم 35 / ج غ / ك ع إ ت ع ت / 2026
غليزان في: 05/06/2026

مستخرج من محضر اجتماع المجلس العلمي للكلية

- بمقتضى المرسوم التنفيذي رقم 03-279 المؤرخ في 24 جمادى الثانية 1424 الموافق 23 غشت 2003، المعدل والمتمم، الذي يحدد مهام الجامعة والقواعد الخاصة بتنظيمها وسيرها؛
- وبمقتضى المرسوم التنفيذي رقم 08-130 المؤرخ في 27 ربيع الآخر 1429 الموافق 3 ماي 2008، المتضمن القانون الأساسي للأستاذ الباحث؛
- وبمقتضى المرسوم التنفيذي رقم 20-339 المؤرخ في 06 ربيع الثاني 1442 الموافق 22 نوفمبر 2020، والمتضمن إنشاء جامعة غليزان؛
- وعملا بأحكام المرسوم التنفيذي رقم 24-103 المؤرخ في 26 شعبان 1445 الموافق 7 مارس 2024، المعدل والمتمم للمرسوم التنفيذي رقم 08-130
- وبمقتضى القرار رقم 498 المؤرخ في 04 ماي 2025، الذي يحدد القائمة الاسمية لأعضاء المجلس العلمي لكلية العلوم الاقتصادية والتجارية وعلوم التسيير بجامعة غليزان؛ والمتضمن الأستاذ: عامري رضوان " أستاذ التعليم العالي"، كرئيس للمجلس العلمي للكلية؛
- وبناء على محضر اجتماع المجلس العلمي للكلية المنعقد بتاريخ 24 ماي 2026 في دورته الاستثنائية الأولى للموسم الجامعي 2025_2026

قرر المجلس العلمي للكلية

المصادقة على المطبوعة البيداغوجية الموسومة بـ " الإحصاء 4"، من إعداد الدكتور مزواغي جيلالي، أستاذ محاضر صنف (أ)، موجهة لطلبة السنة الثانية ليسانس شعبة علوم التسيير، واعتمادها مرجعا بيداغوجيا أساسيا لطلبة الكلية

رئيس المجلس العلمي للكلية
أ. عامري رضوان
رئيس المجلس العلمي للكلية

تعتبر نظرية المعاينة والإحصاء الاستدلالي من الركائز الأساسية في المنهجية العلمية الحديثة، حيث تمكن الباحث من استخلاص استنتاجات دقيقة حول مجتمع كامل انطلاقاً من دراسة عينة صغيرة. تهدف هذه المطبوعة إلى تزويد طلبة السنة الثانية ليسانس في ميدان العلوم الاقتصادية، التجارية وعلوم التسيير، بأدوات نظرية وتطبيقية لفهم توزيعات المعاينة، التقدير النقطي والمجالي، واختبار الفرضيات الإحصائية. تشمل المطبوعة ثلاثة عشر درساً موزعة على خمسة فصول، بدءاً من المفاهيم التأسيسية للمعاينة، مروراً بخصائص المقدرات وطرق اشتقاقها (العزوم والمعقولية العظمى)، وصولاً إلى متباينة كرامير-راو واختبارات الفرضيات وفق منهجية نايمان-بيرسون، تعتمد المطبوعة على أمثلة تطبيقية وتمارين محلولة لتعزيز الفهم، وتتطلب من القارئ إلماماً أولياً بمبادئ الاحتمالات والإحصاء الوصفي.



الفهرس العام

الصفحة	المحتويات
	الفصل الأول: نظرية المعاينة والتوزيعات المعاينية
1	الدرس 1 : المفاهيم الأساسية (المجتمع، العينة، العينة العشوائية، أنواع المعاينة، السحب بالإرجاع وبدون إرجاع، العينة العشوائية والإحصاءة)
12	الدرس 2 : توزيع المعاينة للمتوسط وتوزيع المعاينة للنسبة
41	الدرس 3 : توزيع المعاينة للفروق (بين متوسطين ونسبتين) وتوزيع المعاينة للتباين
	الفصل الثاني: نظرية التقدير
63	الدرس 4 : المفاهيم الأساسية للتقدير وطرق إيجاد المقدرات (العزوم، المعقولية العظمى)
72	الدرس 5 : خصائص المقدرات النقطية – للعينات الصغيرة والكبيرة.
79	الدرس 6 : معلومة فيشر ومتباينة كرامير – راو
87	الدرس 7 : التقدير بفواصل الثقة (مجال الثقة للمتوسط والنسبة) في مجتمع واحد.
102	الدرس 8 : التقدير بفواصل الثقة للتباين ولنسبة تباين مجتمعين.
110	الدرس 9 : التقدير بفواصل الثقة للفرق بين متوسطي مجتمعين والفرق بين نسبتيين.
	الفصل الثالث: اختبار الفرضيات الإحصائية
122	الدرس 10 : نظرية اختبار الفرضيات – منهجية نايمن – بيرسون
127	الدرس 11 : الاختبارات المعلمية لعينة واحدة
133	الدرس 12 : اختبارات الفرضيات لمقارنة عينتين (اختبار الفرق بين متوسطين – واختبار الفرق بين نسبتيين)
137	الدرس 13 : اختبار الفرضيات للتباين
143	الدرس 14: بعض الاختبارات اللامعلمية (Test de Kolmogorov-Smirnov, Le test de Cramer -von Mises)
151	الدرس 15: بعض الاختبارات اللامعلمية (Test de Wilcoxon - Mann - Whitney - Test de Kruskal-Wallis)

الفصل الاول نظرية المعاينة وتوزيعاتها

الدرس الأول: نظرية المعاينة – المفاهيم الإحصائية والرياضية

المقدمة التعريفية:

تعتبر نظرية المعاينة حجر الزاوية في الإحصاء الاستدلالي، حيث تمكن الباحث من دراسة جزء صغير من المجتمع (العينة) ثم تعميم النتائج على المجتمع بأكمله بدرجة معينة من الثقة، في هذا الدرس الأول، سننتقل من المستوى النظري إلى التطبيقي عبر سلسلة من الخطوات المنهجية، ستميز بين المجتمع والعينة والإحصاء والمعلمة، سنصنف أنواع العينات العشوائية المختلفة، سنطبق قواعد التوافق لحساب عدد العينات الممكنة، سنحلل معالم المجتمع مثل التباين والوسط الحسابي، سنقيم أهمية العينة العشوائية في تجنب الانحياز، وأخيراً سنصمم عينة بسيطة من واقع بيانات حقيقية. هذا التدرج سيساعدكم على بناء فهم عميق يربط النظرية بالتطبيق.

الهدف العام للدرس

أن يتقن الطالب بنهاية هذا الدرس تطبيق المفاهيم الأساسية لنظرية المعاينة (العينة العشوائية، الإحصاء، معالم المجتمع) على مجتمعات إحصائية حقيقية أو افتراضية، وذلك من خلال التمييز بين أنواع العينات، وحساب عدد العينات الممكنة، وتقدير معالم المجتمع من العينة، وتقييم مدى ملائمة العينة العشوائية للاستدلال الإحصائي.

الأهداف التفصيلية للدرس

- ❖ تعريف المفاهيم الأساسية للمجتمع، العينة، العينة العشوائية، الإحصاء، المعلمة، والرموز الإحصائية المرتبطة بها؛
- ❖ التمييز بين طرق المعاينة من خلال شرح الفرق بين السحب بالإرجاع وبدون إرجاع، وحساب عدد العينات الممكنة في كل حالة؛
- ❖ تطبيق القواعد الرياضية لحساب عدد العينات (التوافق) ومعالم المجتمع (المتوسط، التباين، الانحراف المعياري)؛
- ❖ تحليل وتقييم العينات من خلال تفسير جداول البيانات، وتبرير اختيار نوع العينة المناسب حسب طبيعة المجتمع؛
- ❖ تصميم خطة عينة عشوائية بسيطة لمجتمع حقيقي مع تحديد طريقة السحب وحجم العينة المناسب.

مواضيع الدرس

- ❖ المفاهيم الأساسية؛
- ❖ أنواع العينات العشوائية وطرق السحب
- ❖ الإحصاء (Statistic) مقابل المعلمة (Parameter)؛
- ❖ معالم المجتمع (Population Parameters) ؛
- ❖ المتغيرات العشوائية ودوال الكثافة.

الوقت المتوقع: 4.5 ساعات بين 3 ساعات محاضرات و 1 سا و 30 حصة أعمال موجهة

تمهيد:

تُعتبر العينة العشوائية (Random Sample) والإحصاء (Statistic) مفهومين محوريين في نظرية المعاينة والإحصاء الاستدلالي، فالإحصاء الاستدلالي يهدف إلى استنتاج خصائص مجتمع كامل (Population) بناءً على دراسة جزء منه فقط، وهو ما يُعرف بالعينة، لكن لا يمكن الوثوق بأي عينة، بل يجب أن تكون العينة عشوائية، وأن تُستخلص منها إحصاءات ذات معنى، في هذه المحاضرة، سنتعرف على هذين المفهومين بالتفصيل، مع التركيز على تعريفاتهما الرياضية، خصائصهما، أنواعهما، وكيفية استخدامهما في الاستدلال الإحصائي.

1-المجتمع:

يمكن فهم مصطلح المجتمع عن طريق الأمثلة التالية:

المثال 1: قصد البحث عن نسبة الطلبة المتقنين للغة الانجليزية في كلية العلوم الاقتصادية، التجارية وعلوم التسيير، بجامعة غليزان، يمكننا استجواب العدد الكلي للطلبة والذي يبلغ : 2000 طالب وطالبة.
-المجتمع المدروس هو : اتقان اللغة الانجليزية من طرف طلبة الكلية.
المثال 2: يسعى الباحث إلى معرفة نسبة الأشخاص الذين يستخدمون الهاتف الذكي في الشراء عبر الانترنت، في الجزائر، وبالتالي فإن :

-المجتمع المدروس هو : استخدام الهاتف الذكي في الشراء عبر الانترنت في دولة الجزائر.
المثال 3: عند الرغبة في قياس معدلات النجاح في كلية الاقتصاد، فإن مجتمع الدراسة هو جميع معدلات طلبة الكلية.

المثال 4: يمثل دراسة أطوال التلاميذ في الابتدائي مجتمعا.
إذن يمكن القول أن المجتمع هو عبارة عن مجموعة من القياسات أو القيم أو المشاهدات التي تشير إلى الظاهرة المدروسة، أكثر دلالة من المفردات (أشياء أو أشخاص). وقد يكون المجتمع محدودا أو غير محدود (بن سبع، 2020، صفحة 3).

يرمز لحجم المجتمع بالرمز: N

2- العينة: (أو المعاينة):

المعاينة هي جزء من المجتمع المدروس مختارة على أساس ما، فغالبا ما يصادف الباحث صعوبة في دراسة المجتمع ككل، فإنه يلجأ إلى دراسة عينة منه فقط وتعميم النتائج المتحصل عليها على المجتمع ككل، وهذا ما نهتم به من خلال دراسة مقياس الإحصاء 4.
يرمز لحجم العينة بالرمز n.

العينة العشوائية (Random Sample)

العينة العشوائية هي مجموعة من المفردات المختارة من مجتمع إحصائي بحيث يكون لكل فرد في المجتمع فرصة متساوية ومعروفة في الاختيار، وتكون اختيارات الأفراد مستقلة عن بعضها البعض (في حالة المجتمع اللانهائي أو السحب بالإرجاع) (عبد السلام ، 2026 ، صفحة 76).

نقول أن المتغيرات العشوائية $(x_1, x_2, x_3, \dots, x_n)$ تشكل عينة عشوائية بسيطة (Simple Random Sample – SRS) من مجتمع له دالة توزيع F إذا تحقق الشرطان التاليان:

- الاستقلال (Independence): المتغيرات $(x_1, x_2, x_3, \dots, x_n)$ مستقلة إحصائياً، أي أن قيمة أي مفردة لا تؤثر على قيمة مفردة أخرى.

- التطابق في التوزيع (Identically Distributed): لكل (x_i) نفس دالة التوزيع F الخاصة بالمجتمع الأصلي.

يمكن تصنيف العينات العشوائية إلى:

- العينة العشوائية البسيطة (Simple Random Sample)

هي العينة التي تُعطي لكل مجموعة فرعية من المجتمع بنفس الحجم n نفس احتمالية الاختيار، تُسحب عادة باستخدام أرقام عشوائية أو برامج حاسوبية.

- العينة العشوائية المنتظمة (Systematic Random Sample)

حيث نختار نقطة بداية عشوائية ثم نختار كل k -th فرد من قائمة المجتمع، سهولة التطبيق لكنها قد تكون أقل دقة إذا كان هناك نمط دوري في البيانات.

- العينة العشوائية الطباقية (Stratified Random Sample)

نقسم المجتمع إلى طبقات متجانسة (حسب الجنس، العمر، المنطقة)، ثم نأخذ عينة عشوائية بسيطة من كل طبقة. هذا يقلل خطأ المعاينة ويحسن الدقة. (عبد السلام ، 2026 ، صفحة 79).

- العينة العشوائية العنقودية (Cluster Random Sample)

نقسم المجتمع إلى عناقيد (مثل المدارس، الأحياء)، ثم نختار عينة عشوائية من العناقيد، ونجري مسحاً شاملاً داخل كل عنقود مختار. (عاروري، 2013، الصفحات 16-17).

السحب بالإرجاع وبدون إرجاع

السحب بالإرجاع **Sampling with replacement** (العينة غير النفاذية) بعد اختيار فرد، يُعاد إلى المجتمع قبل الاختيار التالي، وهذا يضمن الاستقلال التام بين المشاهدات.

$$\text{عدد الحالات الممكنة} = N^n$$

السحب بدون إرجاع **Sampling without replacement** (العينة النفاذية) الفرد المختار لا يُعاد إلى المجتمع، هذا يؤدي إلى اعتماد بسيط بين المشاهدات، ولكن عندما يكون المجتمع كبيراً جداً مقارنة بالعينة ($n/N \leq 0.05$) يمكن تجاهل هذا الاعتماد.

$$C_N^n = \text{عدد الحالات الممكنة}$$

العينة العشوائية هي الأساس الذي يُبنى عليه الاستدلال الإحصائي للأسباب التالية:
عدم التحيز (Unbiasedness): باستخدام العينة العشوائية، يمكننا بناء مقدرات (Estimators) قيمتها المتوقعة تساوي قيمة المعلمة الحقيقية للمجتمع.

قياس دقة التقدير: يمكننا حساب الخطأ المعياري (Standard Error) للتقديرات بناءً على تباين العينة العشوائية.

تعميم النتائج: فقط العينة العشوائية تسمح لنا بتعميم نتائج العينة على المجتمع الأصلي بقوانين الاحتمالات.
تجنب الانحياز (Bias): العينات غير العشوائية (مثل عينة الملاءمة أو التطوعية) تؤدي إلى نتائج مضللة وغير قابلة للتعميم (عاروري، 2013، صفحة 76).

مثال توضيحي: إذا أردنا معرفة متوسط دخل سكان مدينة، وأخذنا عينة من الأشخاص الذين نلتقيهم في فضاء تجاري فاخر، فستكون العينة متحيزة للأغنياء، أما العينة العشوائية البسيطة فتعطي كل فرد نفس فرصة الظهور، وبالتالي يكون متوسط العينة مقدراً جيداً للمتوسط الحقيقي.

2- الإحصاءة (Statistic)

الإحصاءة (Statistic) هي قيمة عددية محسوبة تُطبق على بيانات العينة العشوائية ($X_1, X_2, X_3, \dots, X_n$)، ولا تعتمد على أي معالم مجهولة للمجتمع.

بعبارة أخرى: إذا كانت ($X_1, X_2, X_3, \dots, X_n$) هي العينة، فإن أي دالة قابلة للقياس

$$T = f(X_1, \dots, X_n)$$

تُعتبر إحصاءة.

ملاحظة هامة: الإحصاءة هي متغير عشوائي (Random Variable) لأن قيمته تتغير بتغير العينة، له توزيع احتمالي يُسمى التوزيع المعاينة (Sampling Distribution).

الجدول 1: الفرق بين الإحصاء والمعلمة

المعلمة (Parameter)	الإحصاءة (Statistic)	الخاصية
قيمة تصف المجتمع بأكمله	دالة في بيانات العينة	التعريف
حروف إغريقية (μ, σ, P)	حروف لاتينية $(\bar{x}, \bar{p}, \hat{s})$	الرمز
مجهولة عادة	معروفة بعد سحب العينة	القيمة
ثابت (في نفس المجتمع)	متغير عشوائي (يتغير من عينة لأخرى)	هل هو ثابت في نفس المجتمع؟

المصدر: من اعداد الباحث.

المتغيرات العشوائية: random variables

هو متغير تتحدد قيمته على أساس نتيجة تجربة معينة، وترتبط هذه القيم باحتمال الحدث (هاري و ولاس ، 2001 ، صفحة 15)، كرمي قطعة نقدية متجانسة ذات وجهين (P) أو (F) حيث تتمثل قيمة المتغير العشوائي في الواحد الصحيح 1 عند ظهور الوجه (F) وقيمته 0 عند ظهور الكتابة (P).

مثال تطبيقي 1: ليكن لدينا مجتمع إحصائي يمثل معدلات لسة طلبة جامعيين كما يلي:

اسم الطالب	A	B	C	D	E	F
المعدل السنوي	9.5	12	12.5	11.5	13.5	13

- حدد مجموعة قيم المتغير العشوائي x والتي تمثل المعدلات السنوية.

- ما هو عدد الحالات التي ممكن أن نتحصل عليها من خلال سحب عينة من طالبيين (02) بدون إرجاع؟

الحل:

قيم المتغير العشوائي x 9.5 12 12.5 11.5 13.5 13

حجم المجتمع N=6

حجم العينة n=2

بما أن السحب بدون إرجاع، فإن العينة نفاذية وبالتالي فإن عدد الحالات الممكنة:

عدد الحالات الممكنة = C_N^n ومنه: عدد الحالات الممكنة = C_6^2 تذكير: $6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1$

$$C_N^n = \frac{N!}{n! (N - n)!} = \frac{6!}{2! (6 - 2)!} = \frac{6 \times 5 \times 4!}{2! (4)!} = 15$$

عدد الحالات الممكنة هو = 15 حالة

يمكن استخراج هذه الحالات يدويا كما يلي:

(A,B)	(A,C)	(A,D)	(A,E)	(A,F)
(B,C)	(B,D)	(B,E)	(B,F)	(C,D)
(C,E)	(C,F)	(D,E)	(D,F)	(E,F)

يمكن القول بأن: كل ثنائية تمثل عينة ممكنة.

3-معالم المجتمع

يقصد بمعالم المجتمع المؤشرات الإحصائية له، كالمتوسط الحسابي والانحراف المعياري، والتباين،... الخ. كما أن طبيعة التوزيع الاحتمالي للمجتمع تمثل أحد خصائصه (تمار ، 2020 ، صفحة 27).

دالة الكثافة الاحتمالية probability function

هي الدالة تأخذ شكل معادلة أو جدول والتي تعطي كل احتمالات التي يأخذ فيها المتغير العشوائي كل قيمة من القيم الممكنة له (هاري و ولاس ، 2001 ، صفحة 16)، فإذا كان المتغير العشوائي X هو ظهور الكتابة عند رمي قطعة نقدية، بقيمة عند الظهور وقيمة 0 عند عدم الظهور، فإن دالة الكثافة الاحتمالية تكون على الشكل:

$$f(x) = \frac{1}{2}$$

التوقع الرياضي (expectation)

التوقع الرياضي ($E(X)$) للمتغير العشوائي ($X_1, X_2, X_3, \dots, X_n$) والذي دالته الاحتمالية $f(x)$ ، هو القيمة المتوسطة المرجحة لكل القيم الممكنة للمتغير العشوائي X . حيث (هاري و ولاس ، 2001 ، صفحة 19):

$$E(x) = x_1 f(x_1) + x_2 f(x_2) + \dots + x_n f(x_n)$$

الوسط الحسابي (mean)

هو مقياس نزعة مركزية للمتغير العشوائي (X)، فإذا ما كررنا التجربة عددا من المرات فإن الوسط الحسابي يكون القيمة المتوقعة في المتوسط على مدى كل التجارب (هاري و ولاس ، 2001 ، صفحة 20)، ونرمز له بالرمز (μ) .

يتميز الوسط الحسابي باعتماده على جميع المشاهدات، وسهولة فهمه وحسابه، ولا يتأثر باختلاف العينات في المجتمع، ويعتبر أفضل مقياس نزعة مركزية (نبيل جمعة صالح النجار، 2015 ، صفحة 21).

التباين (Variance) والانحراف المعياري (Standard deviation)

هما مقياسان لتشتت قيم المتغير العشوائي، وبعدها عن المتوسط الحسابي، أو انحراف قيم (X) عن متوسطاتها.

المتوسط الحسابي في المجتمع

$$\mu = \frac{\sum x_i}{N}$$

التباين في المجتمع

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N}$$

الانحراف المعياري في المجتمع وهو الجذر التربيعي للتباين

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$$

ففي المثال السابق مثلاً يمكننا حساب معالم المجتمع كما يلي:

- أولاً نرسم للمعدلات السنوية بالرمز (x_i)

- ثانياً نقوم بحساب المتوسط الحسابي كما يلي:

$$\mu = \frac{\sum x_i}{N} = \frac{(10 + 12 + 12.5 + 11.5 + 10.5 + 12)}{6} = \frac{72}{6} = 12$$

$$\mu = 12$$

- ثالثاً نقوم بحساب كل ثنائية $(x_i - \mu)^2$

اسم الطالب	المعدل السنوي (x_i)	$(x_i - \mu)$	$(x_i - \mu)^2$
A	9.5	$9.5 - 12 = (-2.5)$	6.25
B	12	$12 - 12 = (0)$	0
C	12.5	$12.5 - 12 = (0.5)$	0.25
D	11.5	$11.5 - 12 = (-0.5)$	0.25
E	13.5	$13.5 - 12 = (1.5)$	2.25
F	13	$13 - 12 = (1)$	1
Σ	72	////////	10

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N} = \frac{10}{6} = 1.667$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{1.667} = 1.291$$

مصطلحات الدرس باللغة الانجليزية

Random sample	العينة العشوائية	Population	المجتمع
Population parameters	معالم المجتمع	sample	العينة
Mean	المتوسط الحسابي	sampling	المعاينة
Standard deviation	الانحراف المعياري	Statistical inference	الإحصاء الاستدلالي
Variance	التباين	Sampling with replacement	عينة مسحوبة بالإرجاع
		Sampling without replacement	عينة مسحوبة بدون إرجاع

التقييم الذاتي للدرس: نظرية المعاينة (المفاهيم الأساسية)

أولاً: أسئلة للفهم

السؤال 1 : ما الفرق بين المجتمع (Population) والعينة (Sample) ؟

السؤال 2 : متى نقول عن عينة ما إنها "عينة عشوائية بسيطة؟

السؤال 3 : ما الفرق بين الإحصاء (Statistic) والمعلمة (Parameter) من حيث الرمز والثبات؟

السؤال 4 : لماذا نفضل استخدام العينة العشوائية بدلاً من العينة المتحيزة (مثل عينة الملاءمة)؟

ثانياً: أسئلة تحليلية

السؤال 1 : باحث يريد دراسة متوسط دخل الأسر في مدينة جزائرية، قام باختيار عينة من 100 شخص أمام البنوك الكبرى، هل تعتبر هذه العينة عشوائية؟ علل إجابتك. ما نوع الانحياز (Bias) الذي قد يحدث؟

السؤال 2 : لديك مجتمع حجمه $N=1000$ ، وسحبت منه عينة حجمها $n=100$ بدون إرجاع. هل يمكنك تجاهل اعتماد المشاهدات؟ ولماذا؟

السؤال 3: في مثال الطلاب السنة (المعدلات: 9.5، 12، 12.5، 11.5، 13.5، 13)، قمنا بحساب $\mu=12$ و $\sigma^2=1.667$ لو اخترنا عينة عشوائية من طالبين فقط، هل تتوقع أن يكون متوسط العينة $\bar{x}$ مساوياً تماماً لـ 12 دائماً؟ ولماذا؟

ثالثاً: تطبيق

التمرين التطبيقي 1:

لديك مجتمع إحصائي يتكون من القيم التالية {4,8,12,16} : {4,8,12,16} :

المطلوب:

1. احسب حجم المجتمع N
2. احسب عدد العينات الممكنة إذا قمنا بسحب عينة حجمها $n=2$ بالإرجاع.
3. احسب عدد العينات الممكنة إذا قمنا بسحب عينة حجمها $n=2$ بدون إرجاع.
4. احسب متوسط المجتمع μ
5. احسب تباين المجتمع σ^2 والانحراف المعياري σ

التمرين التطبيقي 2:

لديك العينة التالية لأعمار 5 أطفال (بالسنوات) $\{3,5,7,5,10\}$:

المطلوب:

1. احسب متوسط العينة $\bar{x}$
2. احسب تباين العينة (S^2) باستخدام المقام $(n-1)$ قارنه مع تباين المجتمع σ^2 باستخدام المقام (n) ما الفرق؟

التمرين التطبيقي 3:

مجتمع من 3 أفراد أعمارهم: 20، 25، 30 سنة.

المطلوب:

1. احسب μ و σ^2 للمجتمع.
2. اكتب جميع العينات الممكنة (بدون إرجاع) حجمها $n=2$
3. لكل عينة، احسب متوسط العينة $\bar{x}$
4. احسب متوسط قيمة $\bar{x}$ عبر جميع العينات. هل يساوي μ ؟ (للتحقق من خاصية عدم التحيز).
5. احسب تباين توزيع المعاينة لـ $\bar{x}$ وقارنه بـ $\frac{\sigma^2}{n} \times \frac{N-n}{N-1}$

رابعاً: تقييم ذاتي سريع

ضع دائرة حول الإجابة الصحيحة:

❖ أي من التالي يُعتبر إحصاءة (Statistic) ؟

A. متوسط المجتمع μ

B. تباين المجتمع σ^2

C. متوسط العينة $\bar{x}$

D. حجم المجتمع N

❖ إذا كانت $n/N \leq 0.05$ في السحب بدون إرجاع، فإننا:

A. نستخدم معادلة N^n

B. نضرب في معامل تصحيح المجتمع المحدود

C. يمكننا تجاهل الاعتماد

D. نضاعف حجم العينة

❖ عدد العينات الممكنة عند سحب عينة حجمها $n=3$ من مجتمع حجمه $N=5$ بدون إرجاع هو:

A. 125

B. 60

C. 15

D. 10

❖ أي من أنواع العينات التالية يقلل خطأ المعاينة عن طريق تقسيم المجتمع إلى طبقات متجانسة؟

A. العينة العشوائية البسيطة

B. العينة العشوائية المنتظمة

C. العينة العشوائية العنقودية

D. العينة العشوائية الطبقيّة

الفصل الاول نظرية المعاينة وتوزيعاتها

الدرس الثاني: توزيع المعاينة للمتوسط وتوزيع المعاينة للنسبة

المقدمة التعريفية:

في هذا الدرس الثاني، سننتقل من المفاهيم التأسيسية إلى التطبيقات العملية عبر سلسلة من الخطوات المنهجية. سنُميّز بين توزيع المعاينة للمتوسط وتوزيع المعاينة للنسبة، سنُحسب متوسط وتباين كل منهما في حالات السحب بالإرجاع وبدون إرجاع، سنُطبق قانون نظرية النهاية المركزية (CLT) لحساب الاحتمالات المرتبطة بمتوسطات ونسب العينات، سنُحلل الحالات التي يستوجب فيها استخدام معامل تصحيح المجتمع المحدود (FPC)، وأخيرًا سنُصمم مسائل تطبيقية لحساب الاحتمالات باستخدام التوزيع الطبيعي وتوزيع ستودنت t هذا التدرج سيساعدكم على بناء قدرة استدلالية قوية تمكنكم من تحليل البيانات الحقيقية واتخاذ قرارات علمية سليمة.

الهدف العام للدرس

أن يُتقن الطالب بنهاية هذا الدرس تطبيق مفاهيم توزيع المعاينة للمتوسط وتوزيع المعاينة للنسبة على مجتمعات إحصائية حقيقية أو افتراضية، وذلك من خلال التمييز بين حالات السحب بالإرجاع وبدون إرجاع، وحساب متوسط وتباين توزيع المعاينة، وتطبيق نظرية النهاية المركزية لحساب الاحتمالات، وتقييم مدى ملائمة استخدام التوزيع الطبيعي أو توزيع ستودنت حسب معطيات المسألة.

الأهداف التفصيلية للدرس

- ❖ يُعرّف الطالب المفاهيم التالية: توزيع المعاينة للمتوسط، توزيع المعاينة للنسبة؛
- ❖ يُطبق الطالب حساب الاحتمالات في توزيع المعاينة للمتوسط، توزيع المعاينة للنسبة؛
- ❖ يُقارن الطالب بين استخدام التوزيع الطبيعي (Z) وتوزيع ستودنت (t) عند حساب الاحتمالات؛
- ❖ يُصمم الطالب نماذج تطبيقية لحساب احتمالات توزيع المعاينة من واقع بيانات.

مواضيع الدرس

- ❖ توزيع المعاينة للمتوسط (Sampling Distribution of the Mean) ؛
- ❖ طبيعة توزيع المعاينة ونظرية النهاية المركزية؛
- ❖ توزيع المعاينة للنسبة (Sampling Distribution of Proportion) ؛
- ❖ تطبيقات عملية وأمثلة محلولة

الوقت المتوقع: 4.5 ساعات بين 3 ساعات محاضرات و 1 ساعة و 30 دقيقة أعمال موجهة

تمهيد:

عندما نقوم بسحب عينة عشوائية واحدة من مجتمع ما، فإننا نحصل على قيمة واحدة لمتوسط العينة ($\bar{X}$) لكن هل يمكننا الوثوق بهذه القيمة الواحدة كتقدير لمتوسط المجتمع (μ)؟ الإجابة هي أننا بحاجة إلى فهم سلوك متوسط العينة عبر عينات متعددة. هذا هو بالضبط ما يقدمه لنا توزيع المعاينة للمتوسط الحسابي (Sampling Distribution of the Sample Mean).

1-توزيع المعاينة للمتوسط:

إذا أخذنا عينات متكررة من مجتمع ما، وقمنا بقياس متوسط لكل عينة ($\bar{x}_i$)، فإننا نجد أن معظم هذه المتوسطات تختلف عن بعضها البعض، ويسمى التوزيع الاحتمالي لمتوسطات العينات بـ "توزيع المعاينة للمتوسط الحسابي". ولكن توزيع المعاينة للمتوسط له أيضا متوسط، يعبر عنه بالرمز $\mu_{\bar{x}}$ ، وانحراف معياري $\sigma_{\bar{x}}$. (خليفة، 2019، صفحة 4). السبب الأساسي لاختلاف متوسطات العينات هو التقلبات العشوائية المتأصلة في عملية المعاينة، كل عينة تحتوي على أفراد مختلفين من المجتمع، وحتى لو كان المجتمع متجانساً نسبياً، فإن اختيار أفراد مختلفين سيؤدي إلى قيم مختلفة قليلاً للمتوسط.

مثال توضيحي بسيط:

لنفرض أن لدينا مجتمعاً صغيراً جداً مكوناً من 5 قيم فقط: {2,4,6,8,10} متوسط المجتمع

$$\mu = 6$$

إذا سحبنا عينة عشوائية حجمها $n=2$ بدون إرجاع، فإننا نحصل على نتائج مختلفة:

العينة	القيم	متوسط العينة
1	(2,4)	3
2	(2,6)	4
3	(2,8)	5
4	(2,10)	6
5	(4,6)	5
6	(4,8)	6
7	(4,10)	7
8	(6,8)	7
9	(6,10)	8
10	(8,10)	9

نلاحظ أن متوسطات العينات تراوحت بين 3 و 9، وهي مختلفة عن بعضها البعض وعن متوسط المجتمع (6) في معظم الحالات، هذا الاختلاف هو ما نسميه خطأ المعاينة (Sampling Error)

خطأ المعاينة هو الفرق بين متوسط العينة ($\bar{x}_i$) ومتوسط المجتمع (μ) في المثال أعلاه:

للينة الأولى: خطأ المعاينة = $3 - 6 = -3$

للينة الرابعة: خطأ المعاينة = $0 - 6 = -6$

للينة العاشرة: خطأ المعاينة = $3 - 6 = -3$

ملاحظة هامة: خطأ المعاينة ليس "خطأ" بالمعنى السلبي، بل هو اختلاف طبيعي ومتوقع بسبب العشوائية.

توقع المتغير العشوائي $\bar{x}_i$: هو نفسه الوسط الحسابي للمجتمع وهو مساو للوسط الحسابي لتوزيع المعاينة (الهوبي، 2014، صفحة 33).

$$E(\bar{x}_i) = \mu_{\bar{x}} = \mu$$

البرهان:

نفترض أن لدينا:

مجتمعاً إحصائياً له متوسط μ محدود.

عينة عشوائية بسيطة $X_1, X_2, \dots, X_n$ مكونة من n مشاهدة مسحوبة من هذا المجتمع.

المتغيرات العشوائية $X_1, X_2, \dots, X_n$ هي مستقلة.

القيمة المتوقعة لكل مشاهدة هي μ . $E[X_i] = \mu$

نبدأ من تعريف متوسط العينة:

$$\bar{X} = (X_1 + X_2 + \dots + X_n) / n$$

بأخذ التوقع الرياضي للطرفين:

$$E[\bar{X}] = E[(X_1 + X_2 + \dots + X_n) / n]$$

بسبب خطية التوقع الرياضي، يمكن إخراج الثابت $n/1$ خارج التوقع:

$$E[\bar{X}] = (1/n) * E[X_1 + X_2 + \dots + X_n]$$

وبسبب خطية التوقع أيضاً، فإن توقع المجموع يساوي مجموع التوقعات:

$$E[\bar{X}] = (1/n) * (E[X_1] + E[X_2] + \dots + E[X_n])$$

ولكننا نعلم من الفرضيات أن:

$$E[X_1] = E[X_2] = \dots = E[X_n] = \mu.$$

بالتعويض:

$$E[\bar{X}] = (1/n) * (\mu + \mu + \dots + \mu) \dots \dots \text{مرة } n$$

$$E[\bar{X}] = (1/n) * (n * \mu)$$

$$E[\bar{X}] = \mu$$

$$\mu_{\bar{X}} = E[\bar{X}] = \mu$$

حيث $\mu_{\bar{X}}$ هو متوسط توزيع المعاينة لمتوسط العينة.

هذه النظرية تعني أن: على المدى الطويل، و عبر عدد كبير جدًا من العينات، فإن متوسط متوسطات العينات يتطابق مع متوسط المجتمع الأصلي.

فرغم أن أي عينة واحدة قد تعطي تقديرًا أعلى أو أقل من (μ) أي أن $\bar{x}_i$ قد يكون أكبر أو أصغر من (μ) ، فإن هذه التقديرات "تتوازن" حول القيمة الحقيقية μ عندما نأخذ عددًا كبيرًا من العينات. لنفرض أنك ترمي سهمًا نحو هدف، فأحيانًا يصيب السهم يسار الهدف، وأحيانًا يصيب يمينه، وأحيانًا فوقه أو تحته، لكن متوسط مواقع جميع الأسهم سيكون مركز الهدف تمامًا، هنا، مركز الهدف هو μ ، ومواقع الأسهم هي $\bar{x}_i$ المختلفة، ومتوسطها هو $\mu_{\bar{X}}$ متوسط الحسابي لمتغير توزيع المعاينة.

الوسط الحسابي لمتغير توزيع المعاينة

الوسط الحسابي هو معدل المشاهدات في توزيع المعاينة (نبيل جمعة صالح النجار، 2015، صفحة 22)، أي كل معاينة ولها وسطها الحسابي، وبالتالي يصبح لدينا توزيع جديد مستمد من متوسطات المعاينات المسحوبة، حيث يصبح لدينا مجموعة من المتوسطات للمعاينات: $\bar{X}_1, \bar{X}_2, \bar{X}_3, \dots, \bar{X}_i$ ، ويعطى بالعلاقة التالية:

$$\bar{X}_i = \frac{\sum x_i}{n} \quad \text{متوسط العينة } \bar{X}_i$$

$$\mu_{\bar{X}} = \frac{\sum \bar{X}_i}{\text{عدد الحالات الممكنة}} \quad \text{متوسط توزيع المعاينة } \mu_{\bar{X}}$$

يمكن من خلال هذا المثال أن نستطيع فهم توزيع المعاينة للمتوسطات:

اسم الطالب	A	B	C	D	E	F
المعدل السنوي	9.5	12	12.5	11.5	13.5	13

نسحب عينة من طالبين بدون إرجاع، أي $n=2$ وحجم المجتمع $N=6$

بما أن السحب بدون إرجاع فإن العينة نفاذية:

عدد الحالات الممكنة C_N^n ومنه: عدد الحالات الممكنة C_6^2

تذكير: $6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1$

$$C_N^n = \frac{N!}{n!(N-n)!} = \frac{6!}{2!(6-2)!} = \frac{6 \times 5 \times 4!}{2!(4)!} = 15$$

العينات	المتوسطات للعينة	$\bar{X}_i$	
(A,B)	(9.5+12)/2	10.75	$\bar{X}_1$
(A,C)	(9.5+12.5)/2	11	$\bar{X}_2$
(A,D)	(9.5+11.5)/2	10.5	$\bar{X}_3$
(A,E)	(9.5+13.5)/2	11.5	$\bar{X}_4$
(A,F)	(9.5+13)/2	11.25	$\bar{X}_5$
(B,C)	(12+12.5)/2	12.25	$\bar{X}_6$
(B,D)	(12+11.5)/2	11.75	$\bar{X}_7$
(B,E)	(12+13.5)/2	12.75	$\bar{X}_8$
(B,F)	(12+13)/2	12.5	$\bar{X}_9$
(C,D)	(12.5+11.5)/2	12	$\bar{X}_{10}$
(C,E)	(12.5+13.5)/2	13	$\bar{X}_{11}$
(C,F)	(12.5+13)/2	12.75	$\bar{X}_{12}$
(D,E)	(11.5+13.5)/2	12.5	$\bar{X}_{13}$
(D,F)	(11.5+13)/2	12.25	$\bar{X}_{14}$
(E,F)	(13.5+13)/2	13.25	$\bar{X}_{15}$
	Σ	180	

نلاحظ أن كل عينة ولها متوسط حسابي وبالتالي لدينا 15 معاينة أي لدينا 15 متوسط حسابي للعينة.

$$\bar{X}_1, \bar{X}_2, \bar{X}_3, \dots, \bar{X}_{15}$$

إذن أصبحت لدينا مجموعة المتوسطات $\bar{X}_i$ ، ولها هي الأخرى متوسط حسابي نرسم له بـ

$$\mu_{\bar{X}}$$

حيث:

$$\mu_{\bar{x}} = \frac{\sum \bar{X}_i}{\text{عدد الممكنة الحالات}}$$

$$\mu_{\bar{x}} = \frac{180}{15} = 12$$

التباين والانحراف المعياري لمتغير توزيع المعاينة

إذا كان لدينا مجتمعا غير محدود الحجم أو محدود الحجم، وسحبنا منه عينة عشوائية فإن تباين توزيع متوسطات المعاينة يعطى بالعلاقة التالية (الهوبي، 2014، صفحة 33):

أ/ سحب بالإرجاع (عينة غير نفاذية):

$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$	عينة غير نفاذية (سحب بالإرجاع)	تباين توزيع المعاينة $\sigma_{\bar{x}}^2$
$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$	عينة غير نفاذية (سحب بالإرجاع)	الانحراف المعياري لتوزيع المعاينة $\sigma_{\bar{x}}$

ب/ سحب بدون إرجاع (عينة غير نفاذية):

$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right)$	عينة نفاذية (سحب بدون إرجاع)	تباين توزيع المعاينة $\sigma_{\bar{x}}^2$
$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\left(\frac{N-n}{N-1} \right)}$	عينة نفاذية (سحب بدون إرجاع)	الانحراف المعياري لتوزيع المعاينة $\sigma_{\bar{x}}$

تسمى النسبة $\left(\frac{N-n}{N-1} \right)$ معامل الإرجاع.

• من المثال السابق يمكن حساب الانحراف المعياري لتوزيع المعاينة والتباين:

المتوسط الحسابي للمجتمع μ

$$\mu = \frac{\sum x_i}{N} = 12$$

التباين في المجتمع σ^2

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N} = 1.667$$

الانحراف المعياري في المجتمع σ

$$\sigma = \sqrt{\sigma^2} = \sqrt{1.667} = 1.291$$

متوسطات العينة $\bar{X}_i$

هي محسوبة في الجدول: $\bar{X}_i = \frac{\sum x_i}{n}$

متوسط توزيع المعاينة $\mu_{\bar{x}}$

$$\mu_{\bar{x}} = \frac{\sum \bar{X}_i}{\text{عدد الحالات الممكنة}} = 12$$

تباين توزيع المعاينة $\sigma_{\bar{x}}^2$

$$\begin{aligned}\sigma_{\bar{x}}^2 &= \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right) \\ &= \frac{1.667}{2} \left(\frac{6-2}{6-1} \right) = 0.6668\end{aligned}$$

الانحراف المعياري لتوزيع المعاينة $\sigma_{\bar{x}}$

$$\begin{aligned}\sigma_{\bar{x}} &= \frac{\sigma}{\sqrt{n}} \sqrt{\left(\frac{N-n}{N-1} \right)} \\ &= \frac{1.291}{\sqrt{2}} \sqrt{\left(\frac{6-2}{6-1} \right)} \\ &= 0.8165\end{aligned}$$

2- طبيعة توزيع المعاينة

أ/ توزيع المعاينة للمتوسط الحسابي في حالة مجتمع طبيعي وتباينه معلوم

إذا كان المجتمع يتبع التوزيع الطبيعي بمتوسط حسابي (μ) وتباين معلوم (σ^2)، فإن المعاينة تتبع التوزيع الطبيعي ونكتب (تمار ، 2020، صفحة 29):

$$X_i \sim N(\mu, \sigma^2)$$

في حالة السحب بالإرجاع: (كما هو الحال إذا كان معامل الاستقصاء $\frac{n}{N} < 0.05$)

$$\bar{X}_i \sim N(\mu_{\bar{x}}, \sqrt{\frac{\sigma^2}{n}})$$

$$\bar{X}_i \sim N(\mu_{\bar{x}}, \sigma_{\bar{x}}) \text{ أو}$$

في حالة السحب بدون إرجاع: (كما هو الحال إذا كان معامل الاستقصاء $\frac{n}{N} > 0.05$)

$$\bar{X}_i \sim N(\mu_{\bar{x}}, \sqrt{\frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right)})$$

$$\bar{X}_i \sim N(\mu_{\bar{x}}, \sigma_{\bar{x}}) \text{ أو } \bar{X}_i \sim N(\mu_{\bar{x}}, \sigma_{\bar{x}})$$

السحب بالإرجاع	لا نستخدم معامل الإرجاع
السحب بدون إرجاع	$\frac{n}{N} \leq 0.05$ لا نستخدم معامل الإرجاع
السحب بدون إرجاع	$\frac{n}{N} > 0.05$ نستخدم معامل الإرجاع

ب/ نظرية النهاية المركزية

إذا كان المجتمع يتبع التوزيع الطبيعي بمتوسط حسابي (μ) وتباين (σ^2)، وأخذنا عينات عشوائية بحجم أكبر من 30، فإن المعاينة تتبع التوزيع الطبيعي ($\mu = \mu_{\bar{x}}$) وتباين $\frac{\sigma^2}{n}$ (تمار ، 2020، صفحة 35) نقلا عن (الصياد و مصطفى ، الصفحات 51-55).

مثال تطبيقي 1:

إذا كان مجتمع ما يتبع التوزيع الطبيعي حجمه 1000 بمتوسط ($\mu = 30$)، وانحراف معياري ($\sigma = 6$). المطلوب:

- أحسب المتوسط والانحراف المعياري لتوزيع المعاينة للمتوسطات في حالات التالية:

$$n=25$$

$$n=87$$

الحل:

لدينا

$$N=1000$$

$$\mu = 30$$

$$\sigma = 6$$

الحالة الأولى: $n=25$

بما أن:

$$\left\{ \begin{array}{l} \text{الطبيعي التوزيع يتبع المجتمع} \\ n = 25 < 30 \end{array} \right.$$

فإن متغير توزيع المعاينة $\bar{X}_i$ يتبع التوزيع الطبيعي ومنه: $\mu = \mu_{\bar{x}} = 30$

ونحسب معامل الاستقصاء:

$$\frac{n}{N} = \frac{25}{1000} = 0.025 < 0.05$$

وبالتالي لا نستخدم معامل التصحيح:

$$\bar{X}_i \sim N(\mu_{\bar{x}}, \sqrt{\frac{\sigma^2}{n}})$$

$$\bar{X}_i \sim N(\mu_{\bar{x}}, \sigma_{\bar{x}}) \text{ أو}$$

المتوسط الحسابي لتوزيع المعاينة :

$$\mu = \mu_{\bar{x}} = 30$$

التباين لتوزيع المعاينة:

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{6^2}{27} = 1.333$$

الانحراف المعياري لتوزيع المعاينة:

$$\sigma_{\bar{x}} = 1.153$$

$$\bar{X}_i \sim N(30, 1.153)$$

الحالة الثانية: n=87

بما أن:

$$\left\{ \begin{array}{l} \text{الطبيعي التوزيع يتبع المجتمع} \\ n = 58 < 30 \end{array} \right.$$

فإن متغير توزيع المعاينة $\bar{X}_i$ يتبع التوزيع الطبيعي ومنه:

$$\mu = \mu_{\bar{x}} = 30$$

ونحسب معامل الاستقصاء:

$$\frac{n}{N} = \frac{87}{1000} = 0.087 > 0.054$$

وبالتالي:

$$\bar{X}_i \sim N(\mu_{\bar{x}}, \frac{\sigma^2}{n} (\frac{N-n}{N-1}))$$

$$\bar{X}_i \sim N(\mu_{\bar{x}}, \sigma_{\bar{x}}^2) \text{ أو}$$

المتوسط الحسابي لتوزيع المعاينة :

$$\mu = \mu_{\bar{x}} = 30$$

التباين لتوزيع المعاينة:

$$\begin{aligned}\sigma_x^2 &= \frac{\sigma^2}{n} \left(\frac{N-n}{N-1} \right) \\ &= \frac{6^2}{58} \left(\frac{1000-87}{1000-1} \right) \\ &= 0.620 \times 0.914 \\ &= 0.56668\end{aligned}$$

الانحراف المعياري لتوزيع المعاينة:

$$\sigma_{\bar{x}} = 0.752$$

$$\bar{X}_i \sim N(30, 0.752)$$

جـ/ حساب الاحتمال في حالة المجتمع الطبيعي وتباينه معلوم (بوالشعور، 2018، صفحة 20):

إذا كان المجتمع يتبع التوزيع الطبيعي بمتوسط حسابي (μ) وتباين (σ)، فإن حساب احتمال $p(X_i)$ يتم حسابه عن طريق المتغير العشوائي المعياري Z الذي يتبع التوزيع الطبيعي بمتوسط حسابي 1 وانحراف معياري 0

$$Z \sim N(1, 0)$$

$$Z = \frac{x - \mu}{\sigma} \text{ حيث:}$$

إذا كان متغير توزيع المعاينة يتبع التوزيع الطبيعي بمتوسط حسابي ($\mu_{\bar{x}}$) وتباين ($\sigma_{\bar{x}}^2$)، فإن حساب احتمال $p(\bar{X}_i)$ يتم حسابه عن طريق المتغير العشوائي المعياري Z حيث:

$$Z = \frac{\bar{x}_i - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

• مثال تطبيقي 2 (نفس المثال السابق):

إذا كان مجتمع ما يتبع التوزيع الطبيعي حجمه 1000 بمتوسط ($\mu = 30$)، وانحراف معياري ($\sigma = 6$).

• أحسب الاحتمالات المتعلقة بالمجتمع في حالة: $p(X \leq 36)$, $p(X > 24)$, $p(X > 31)$,
احسب في الحالتين السابقتين ($n=25$, $n=87$)، الاحتمال عندما يكون: $\bar{X} < 28$ و $\bar{X} > 31$ ثم
 $29 < \bar{X} < 32$
الحل:

• حساب الاحتمالات في مجتمع الدراسة:

$$p(X \leq 36) = P(Z \leq \frac{x - \mu}{\sigma})$$

$$= P(Z \leq \frac{36-30}{6})$$

$$= P(Z \leq 1) = \Phi(1)$$

من جدول توزيع (Z) نجد:

$$\Phi(1) = 0,8413$$

ومنه: $P(X \leq 35) = 0.8413$

$$P(X > 31) = P(Z > \frac{x-\mu}{\sigma})$$

$$= P(Z > \frac{31-30}{6})$$

$$= P(Z > 0.1667)$$

$$= 1 - P(Z \leq 0.1667)$$

$$= 1 - \Phi(0.1667)$$

$$= 1 - 0.5662 = 0.4338$$

• قاعدة: $P(Z > a) = 1 - P(Z \leq a) = 1 - \Phi(a)$

$$p(X > 24) = P\left(Z > \frac{x - \mu}{\sigma}\right)$$

$$= P\left(Z > \frac{24 - 30}{6}\right)$$

$$= P(Z > -1)$$

$$= \Phi(1) = 0.8413$$

• قاعدة: $P(Z > -a) = p(Z \leq a) = \Phi(a)$

• الحالة n=25 بمتوسط حسابي لتوزيع المعاينة $\mu_{\bar{x}} = 30$ وانحراف معياري لتوزيع المعاينة:

$$\sigma_{\bar{x}} = 1.153$$

$$p(\bar{X} < 28) = p\left(Z < \frac{\bar{x}_i - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right)$$

$$= p\left(Z < \frac{28 - 30}{1.153}\right)$$

$$= p(z < -1.73)$$

$$= 1 - \Phi(1.73)$$

• قاعدة: $P(Z \leq -a) = 1 - p(Z \leq a) = 1 - \Phi(a)$

نستخرج قيمة $\Phi(1.73)$ من جدول قانون التوزيع الطبيعي، وهي: 0.9582

$$p(\bar{X} < 28) = 1 - 0.9582 = 0.0418$$

$$p(\bar{X} > 30) = p\left(Z > \frac{\bar{x}_i - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right)$$

$$= p\left(Z > \frac{31 - 30}{1.153}\right) = p(z > 0.86)$$

$$= 1 - \Phi(0.86)$$

$$= 1 - 0.8051$$

$$p(\bar{X} > 31) = 0.1949$$

• قاعدة: $P(\beta \leq Z \leq a) = p(Z \leq a) - p(Z \leq \beta) = \Phi(a) - \Phi(\beta)$

$$p(29 < \bar{X} < 32) = p(\bar{X} < 32) - p(\bar{X} < 29) =$$

$$= p\left(Z < \frac{\bar{x}_1 - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right) - p\left(Z < \frac{\bar{x}_2 - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right)$$

$$= p\left(Z < \frac{32 - 30}{1.153}\right) - p\left(Z < \frac{29 - 30}{1.153}\right)$$

$$= p(Z < 1.73) - p(Z < -0.86)$$

$$= 0.9582 - 0.1949 = 0.7633$$

$$p(29 < \bar{X} < 32) = 0.7633$$

• الحالة n=58 بمتوسط حسابي لتوزيع المعاينة $\mu_{\bar{x}} = 30$ وانحراف معياري لتوزيع المعاينة:

$$\sigma_{\bar{x}} = 0.764$$

$$p(\bar{X} < 28) = p\left(Z < \frac{\bar{x}_i - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right)$$

$$= p\left(Z < \frac{28 - 30}{0.752}\right)$$

$$= p(z < -2.65)$$

$$= 1 - \Phi(2.65)$$

نستخرج قيمة $\Phi(2.65)$ من جدول قانون التوزيع الطبيعي، وهي: **0.9960**

$$p(\bar{X} < 28) = 1 - 0.9960 = 0.004$$

$$p(\bar{X} > 31) = p\left(Z > \frac{\bar{x}_i - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right)$$

$$= p\left(Z > \frac{31 - 30}{0.752}\right)$$

$$\begin{aligned}
&= p(z > 1.33) \\
&= 1 - \Phi(1.33) \\
&= 1 - 0.9082 = 0.0918 \\
&p(\bar{X} > 31) = 0.0918 \\
p(29 < \bar{X} < 32) &= p(\bar{X} < 32) - p(\bar{X} < 29) = \\
&= p\left(Z < \frac{\bar{x}_1 - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right) - p\left(Z < \frac{\bar{x}_2 - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right) \\
&= p\left(Z < \frac{32 - 30}{0.752}\right) - p\left(Z < \frac{29 - 30}{0.752}\right) \\
&= p(Z < 2.66) - p(Z < -1.33) \\
&= 0.9961 - (1 - 0.9082) = 0.9043 \\
p(29 < \bar{X} < 32) &= 0.9043
\end{aligned}$$

مثال تطبيقي 3:

500 بطيخة، لها متوسط وزن 5.02 كغ، بانحراف معياري 0.3 كغ. سحبنا منها 100 بطيخة. - أحسب احتمال أن تكون 100 بطيخة المسحوبة تزن أكثر من 510 كغ.

الحل:

$$N=500$$

$$\mu = 5.02$$

$$\sigma = 0.3$$

بما أن:

$$n = 100 > 30$$

فان:

المعينة تتبع التوزيع الطبيعي ومنه:

$$\mu = \mu_{\bar{x}} = 5.02$$

السحب بدون إرجاع، ونحسب معامل الاستقصاء:

$$\frac{n}{N} = \frac{100}{500} = 0.2 > 0.05$$

وبالتالي نستخدم معامل التصحيح نستخدم معامل التصحيح.

$$\begin{aligned}\bar{X}_i &\sim N\left(\mu_{\bar{x}}, \frac{\sigma^2}{n} \left(\frac{N-n}{N-1}\right)\right) \\ \bar{X}_i &\sim N\left(5.02, \frac{0.3^2}{100} \left(\frac{500-100}{500-1}\right)\right) \\ \bar{X}_i &\sim N\left(5.02, \frac{0.3^2}{100} \left(\frac{500-100}{500-1}\right)\right) \\ \bar{X}_i &\sim N(5.02, 7.21)\end{aligned}$$

المتوسط الحسابي لتوزيع المعاينة :

$$\boxed{\mu = \mu_{\bar{x}} = 5.02}$$

التباين لتوزيع المعاينة:

$$\boxed{\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \left(\frac{N-n}{N-1}\right) = 7.21}$$

الانحراف المعياري لتوزيع المعاينة:

$$\boxed{\sigma_{\bar{x}} = 0.027}$$

100 بطيخة اكبر من 510 كغ → متوسط 100 بطيخة اكبر من 5.1 كغ.

$$\begin{aligned}p(\bar{X} > 5.1) &= p\left(Z > \frac{\bar{x}_i - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right) \\ &= p\left(Z > \frac{5.1 - 5.02}{0.027}\right) \\ &= p(z > 2.96) \\ &= 1 - p(z < 2.96) \\ &= 1 - 0.9985 = 0.0015\end{aligned}$$

ج/ حساب الاحتمال في حالة المجتمع الطبيعي وتباينه مجهول (بوالشعور، 2018، صفحة 20):

إذا كان المجتمع يتبع التوزيع الطبيعي بمتوسط حسابي (μ) وتباين مجهول، فإنه يمكن تعويض هذا الأخير بتباين توزيع العينة S ، ومنه يمكن حساب احتمال $p(X_i)$ يتم حسابه عن طريق المتغير العشوائي المعياري t الذي يتبع توزيع ستودنت بدرجة حرية $v=n$ مساوية لعدد عناصر العينة.

$$t \sim t(v = n)$$

حيث:

$$t = \frac{\bar{x}_i - \mu}{\frac{s}{\sqrt{n}}}$$

مثال تطبيقي

مثال تطبيقي:

نفترض أن لدينا مجتمعاً يتبع التوزيع الطبيعي:

- متوسط المجتمع: $\mu=50$ قيمة مفروضة
- تباين المجتمع σ^2 مجهول

قمنا بسحب عينة عشوائية حجمها $n=16$ وحصلنا على النتائج التالية:

رقم الملاحظة	القيمة
1	48
2	52
3	47
4	51
5	49
6	53
7	46
8	50
9	48
10	52

حساب متوسط العينة

$$\bar{X}_i = \frac{\sum x_i}{n} = 794/16 = 49.625$$

حساب انحراف معياري للعينة (S)

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

$$S = \sqrt{S^2}$$

نحسب الانحرافات :

$(X_i - \bar{X})^2$	$X_i - \bar{X}$	X_i
2.6406	-1.625	48
5.6406	2.375	52
6.8906	-2.625	47
1.8906	1.375	51
0.3906	-0.625	49
11.3906	3.375	53
13.1406	-3.625	46
0.1406	0.375	50
2.6406	-1.625	48
5.6406	2.375	52
6.8906	-2.625	47
1.8906	1.375	51
0.3906	-0.625	49
0.1406	0.375	50
2.6406	-1.625	48
11.3906	3.375	53

$$S^2 = 4.8333$$

$$S = 2.198$$

حساب قيمة t المطلوب: حساب احتمال أن يكون متوسط العينة أقل من 49.625 (أي $P(\bar{X} \leq 49.625)$)

$$\begin{aligned}
 t &= \frac{\bar{x}_i - \mu}{\frac{s}{\sqrt{n}}} \\
 &= \frac{48 - 50}{2.198/\sqrt{16}} \\
 &= -3.64
 \end{aligned}$$

تحديد درجات الحرية

$$v = n - 1 = 16 - 1 = 15$$

إيجاد الاحتمال من جدول توزيع t

القيمة الحرجة عند $v=15$ و $\alpha = 0.001$

$$t_{0.001} = 3.733$$

القيمة الحرجة عند $v=15$ و $\alpha = 0.0005$

$$t_{0.0005} = 4.073$$

بما ان:

$$|t| = 3.64$$

تقع بين القيمتين: $0.0005 < P(T < -3.64) < 0.001$

$$P(\bar{X} < 48) \approx 0.0012$$

عندما يكون التباين مجهولاً، لا يمكننا استخدام التوزيع الطبيعي (Z)، بل نستخدم توزيع ستودنت t مع $v=n-1$ درجة حرية. هذا التوزيع له ذيول أثقل من التوزيع الطبيعي، مما يعكس عدم اليقين الإضافي الناتج عن تقدير التباين من العينة نفسها.

3-توزيع المعاينة للنسبة:

في العديد من الدراسات الإحصائية، لا نهتم بقياس قيمة كمية (مثل الطول أو الوزن أو الدخل)، بل نهتم بتصنيف الأفراد أو العناصر ضمن فئتين متقابلتين، مثل: ناجح/راسب، مؤيد/معارض، ذكر/أنثى، معيب/سليم، مصاب/سليم. في هذه الحالات، يكون المتغير العشوائي من النوع الثنائي (Binary) أو برنولي (Bernoulli)، حيث يأخذ إحدى قيمتين: 1 (نجاح) أو 0 (فشل). الهدف الأساسي في مثل هذه الدراسات هو تقدير نسبة النجاح (p) في المجتمع، وذلك من خلال حساب نسبة النجاح في العينة ($\bar{p}$) هذا الموضوع يعرف باسم توزيع المعاينة للنسب (Sampling Distribution of Proportions) (عاروري، 2013، صفحة 35).

سيركز هذا الدرس على الجوانب النظرية والتطبيقية لتوزيع المعاينة للنسب، بدءاً من المفاهيم الأساسية لمتغير برنولي، مروراً بالتوزيع الدقيق للتوزيع ذي الحدين، وصولاً إلى التقريب الطبيعي القائم على نظرية النهاية المركزية (CLT)، مع تقديم أمثلة تطبيقية وتمارين محلولة.

نسمي توزيع المجتمع بتوزيع برنولي أو ذي الحدين، إذا كان الحدث فيه يحتمل نتيجتين (نجاح/فشل) حيث احتمال النجاح p واحتمال عدم النجاح q علما بأن $p+q=1$. (نبيل جمعة صالح النجار، 2015، صفحة 105)

مثال: عند رمي قطعة نقدية ذات وجه وكتابة، لمرة واحدة فإن احتمال ظهور الصورة مساو لاحتمال ظهور الكتابة، لذا يمكن القول:

- احتمال ظهور الصورة هو $p=50\%$ أو $p=0.5$
- احتمال عدم ظهور الصورة هو $q=50\%$ أو $q=0.5$
- دالة الكتلة الاحتمالية (PMF) لمتغير برنولي:

دالة كتلة الاحتمال $f(x)$ في هذه الحالة هي:

$$f(x) = P(X = x) = \begin{cases} p^x(1-p)^{1-x} & x = 0, 1 \\ 0 & x \neq 0, 1 \end{cases}$$

حيث:

- $P(X = x)$: احتمال أن يكون المتغير العشوائي (X) يأخذ القيمة (x) .
- (p) : احتمال النجاح.
- $(1-p)$: احتمال الفشل ويرمز له كذلك بـ (q) .

نقول بأن معلمة توزيع بيرنولي هي قيمة احتمال النجاح (p) ونرمز لهذا التوزيع بالرمز:

$$X \sim Ber(p)$$

نعتبر مجتمع ذو حجم N ، و N_p هو عدد المحاولات التي تؤدي إلى نجاح الظاهرة (في مثالنا السابق ظهور الصورة) فإن النسبة في المجتمع p تعطى بالعلاقة التالية:

$$p = \frac{N_p}{N}$$

نعتبر أن n حجم العينة المسحوبة من المجتمع، و n_p هو عدد المحاولات التي تؤدي إلى نجاح الظاهرة في العينة فإن النسبة في العينة $\bar{p}$ فان توزيع المعاينة يمكن ان يكون قريبا من التوزيع الطبيعي بوسط حسابي: (نبيل جمعة صالح النجار، 2015، صفحة 105).

$$\mu_{\bar{p}} = p$$

$$\bar{p} = \frac{n_p}{n}$$

وتباين:

$$\sigma_{\bar{p}}^2 = \frac{pq}{n}$$

نظرية النهاية المركزية: إذا كانت $n \geq 30$ فإن توزيع المعاينة يقترب من التوزيع الطبيعي. حيث:

$$\bar{p} \sim N(\mu_{\bar{p}}, \sigma_{\bar{p}})$$

نحاول حساب التوقع الرياضي بناء على العلاقات الرياضية المعروفة، نعلم أن: $\mu = E(x)$

$$\begin{aligned} \mu = E(x) &= \sum_x^i X f(x) \\ &= \sum_{X=0}^{i=1} X f(x) \end{aligned}$$

نقوم بتعويض:

$$\begin{aligned} f(x) &= p^x(1-p)^{1-x} \\ \mu = E(x) &= \sum_{X=0}^{i=1} X (p^x(1-p)^{1-x}) \\ &= 0(p^0(1-p)^{1-0}) + 1(p^1(1-p)^{1-1}) \\ &= 0 + p \\ \boxed{\mu = E(x) = p} \end{aligned}$$

يمكن حساب التباين في المجتمع:

$$\sigma_x^2 = E(x^2) - [E(x)]^2$$

نعلم أن: $E(x) = p$
وبالتالي فإن:

$$[E(x)]^2 = p^2$$

كذلك من جهة أخرى:

$$\begin{aligned} E(x^2) &= p \\ \text{لأن:} \end{aligned}$$

$$\begin{aligned} E(x^2) &= \sum_{X=0}^{i=1} X^2 (p^x(1-p)^{1-x}) \\ &= 0^2(p^0(1-p)^{1-0}) + 1^2(p^1(1-p)^{1-1}) = p \\ \sigma_x^2 &= E(x^2) - [p]^2 \end{aligned}$$

إذن:

$$\sigma_x^2 = E(x^2) - [E(x)]^2$$

$$\sigma_x^2 = p - p^2$$

نقوم بتبسيط العلاقة على أساس p عامل مشترك، فنجد:

$$\sigma_x^2 = p(1 - p)$$

نعلم أن: $(1 - p) = q$

التباين في توزيع بيرنولي يعطى على أساس العلاقة:

$$\sigma_x^2 = pq$$

أما الانحراف المعياري:

$$\sigma_x = \sqrt{pq}$$

باستخدام خاصية خطية التوقع الرياضي، وباعتماد على ما ورد في هاري وولاس (2001، صفحة 19) حول التوقع الرياضي، نحصل على:

$$E(\bar{p}) = E\left(\frac{Y}{n}\right)$$

$$= \frac{1}{n} E(Y)$$

$$= \frac{1}{n} \cdot np$$

$$= p$$

$$\mu_{\bar{p}} = p$$

تباين توزيع المعاينة لـ $\bar{p}$

بحسب ما ورد في بن سبع (2020، صفحة 8)، فإن تباين توزيع المعاينة للنسبة يحسب كما يلي:

$$Var(\bar{p}) = Var\left(\frac{Y}{n}\right)$$

$$= \frac{1}{n^2} Var(Y)$$

$$= \frac{1}{n^2} (np(1 - p))$$

$$= \frac{p(1-p)}{n}$$

$$\sigma_{\bar{p}}^2 = \frac{pq}{n} = \frac{p(1-p)}{n}$$

ومنه الانحراف المعياري لمتغير تويح المعاينة في النسبة:

$$\sigma_{\bar{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{p(1-p)}{n}}$$

$$\mu_{\bar{p}} = p$$

الوسط الحسابي لتوزيع معاينة
النسبة

$$\sigma_{\bar{p}}^2 = \frac{pq}{n} = \frac{p(1-p)}{n}$$

السحب
بالإرجاع

$$\sigma_{\bar{p}}^2 = \frac{pq}{n} \left(\frac{N-n}{N-1} \right)$$

التباين لتوزيع معاينة النسبة

$$= \frac{p(1-p)}{n} \left(\frac{N-n}{N-1} \right)$$

السحب بدون
إرجاع

$$\sigma_{\bar{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{p(1-p)}{n}}$$

السحب
بالإرجاع

$$\sigma_{\bar{p}} = \sqrt{\frac{pq}{n} \left(\frac{N-n}{N-1} \right)}$$

الانحراف المعياري لتوزيع
معاينة النسبة

$$= \sqrt{\frac{p(1-p)}{n} \left(\frac{N-n}{N-1} \right)}$$

السحب بدون
إرجاع

نظرية الاحتمال في توزيع النسبة: سحبت عينة عشوائية حجمها (n) من مجتمع احصائي يخضع للتوزيع

ذو حدين، بوسط حسابي μ وتباين σ^2 فتوزيع المعاينة الى $\bar{p}$ هو قريب من التوزيع الطبيعي بوسط $\mu_{\bar{p}}$

$$p \text{ وانحراف معياري } \sigma_{\bar{p}} = \sqrt{\frac{pq}{n}}$$

لذا فإن القيمة المعيارية لـ Z هي:

$$Z = \frac{\bar{p} - \mu_{\bar{p}}}{\sqrt{\frac{pq}{n}}}$$

$$Z = \frac{\bar{p} - \mu_{\bar{p}}}{\sigma_{\bar{p}}}$$

شروط التقريب الطبيعي:

لكي يكون التقريب الطبيعي مقبولاً، يجب أن تتحقق الشروط التالية (بوالشعور، 2018، صفحة 16):

$$np \geq 5$$

$$n(1 - p) \geq 5$$

أحياناً تُستخدم شروط أكثر تحفظاً :

$$np \geq 10$$

$$n(1 - p) \geq 10$$

مثال تطبيقي 1:

مصنع للمصابيح، أنتج 10000 مصباح من بينها 200 مصباحاً فاسداً لأسباب تقنية.

إذا اشترى تاجر 500 مصباح من هذا المصنع، فأوجد:

- طبيعة توزيع العينة لنسبة المصابيح المعيبة التي اشتراها التاجر.
- ما هو احتمال أن تكون نسبة المصابيح الفاسدة 0.05 على الأقل.

الحل:

$$N=10000$$

$$N_p=200$$

$$p = \frac{N_p}{N} = \frac{200}{10000} = 0.02 \quad \text{alors} \quad q = 0.98$$

بما أن $n > 30$ فإن توزيع المعاينة يقترب من التوزيع الطبيعي حيث:

$$\bar{p} \sim N(\mu_{\bar{p}}, \sigma_{\bar{p}})$$

الوسط الحسابي لتوزيع معاينة النسبة

$$\mu_{\bar{p}} = p = 0.02$$

التباين لتوزيع معاينة النسبة

السحب بدون إرجاع

$$\begin{aligned}\sigma_{\bar{p}}^2 &= \frac{pq}{n} \left(\frac{N-n}{N-1} \right) = \\ \sigma_{\bar{p}}^2 &= \frac{0.02 \times 0.98}{500} \left(\frac{10000 - 500}{10000 - 1} \right) \\ &= 0.00003724\end{aligned}$$

الانحراف المعياري لتوزيع معاينة النسبة

$$\sigma_{\bar{p}} = \sqrt{\frac{pq}{n} \left(\frac{N-n}{N-1} \right)} = \sqrt{0.00003724} = 0.006102$$

$$\bar{p} \sim N(0.02, 0.006102)$$

- حساب الاحتمال:

$$\begin{aligned}p(\bar{p} > 0.05) &= p\left(Z > \frac{\bar{p} - \mu_{\bar{p}}}{\sigma_{\bar{p}}}\right) = p\left(Z > \frac{0.05 - 0.02}{0.006102}\right) = p(z > 4.916) \\ &= 1 - p(z < 4.916) = 1 - 0.9999995 = 0.0000005\end{aligned}$$

هذا الاحتمال الصغير جدًا يشير إلى أن الفرق بين 0.05 و 0.02 (أي 3%) هو فرق كبير جدًا من الناحية الإحصائية عندما يكون الخطأ المعياري 0.0061 فقط. ومن المستبعد جدًا أن يكون هذا الفرق قد حدث بالصدفة.

مثال تطبيقي 2:

تمثل مجموعة طلبة السنة الثانية بكلية العلوم الاقتصادية والتجارية وعلوم التسيير مجتمعًا إحصائيًا، إذ يقدر عددهم بـ 200 طالب، 40% منهم يرغبون في خلق مؤسسة ناشئة (P=0.40)، تم سحب عينة من 50 طالب (سحب بدون إرجاع).

المطلوب:

- ما هو عدد الطلبة الذين يرغبون في خلق مؤسسة ناشئة؟
- ما هي نسبة الطلبة الذين لا يرغبون في خلق مؤسسة ناشئة؟
- ما هو عدد الطلبة الذين لا يرغبون في خلق مؤسسة ناشئة؟
- أوجد متوسط توزيع المعاينة للنسبة $(\mu_{\bar{p}})$.
- أوجد تباين توزيع المعاينة للنسبة $(\sigma_{\bar{p}}^2)$.
- أحسب احتمال أن تكون نسبة العينة $(\bar{p})$ أصغر أو تساوي 42%.
- أحسب احتمال أن تكون نسبة العينة $(\bar{p})$ محصورة ما بين 43% و 45%.

الحل:

$$P=0.40 \quad ; \quad N=200 \quad ; \quad n=50$$

- ما هو عدد الطلبة الذين يرغبون في خلق مؤسسة ناشئة؟

$$PXN=0.4 \times 200 = 80 \text{ طالب}$$

- ما هي نسبة الطلبة الذين لا يرغبون في خلق مؤسسة ناشئة؟

نرمز لها بـ q لأنها تمثل الحدث العكسي:

$$q=1-p=1-0.4=0.6$$

- ما هو عدد الطلبة الذين لا يرغبون في خلق مؤسسة ناشئة؟

$$qXN=0.6 \times 200 = 120$$

عدد الطلبة الذين لا يرغبون في خلق مؤسسة ناشئة هو 120 طالب

- أوجد متوسط توزيع المعاينة للنسبة P ($\mu_{\bar{p}}$).

نظرية النهاية المركزية:

$$n \geq 30$$

فإن توزيع المعاينة يقترب من التوزيع الطبيعي حيث:

$$\bar{p} \sim N(\mu_{\bar{p}}, \sigma_{\bar{p}}^2)$$

$$\mu_{\bar{p}} = p = 0.4$$

- أوجد تباين توزيع المعاينة للنسبة P ($\sigma_{\bar{p}}^2$).

بما أن السحب بدون إرجاع

$$\frac{n}{N} = \frac{50}{200} = 0.25 > 0.05$$

فإننا نستخدم معامل الإرجاع:

$$\begin{aligned} \sigma_{\bar{p}}^2 &= \frac{pq}{n} \left(\frac{N-n}{N-1} \right) \\ &= \frac{p(1-p)}{n} \left(\frac{N-n}{N-1} \right) \\ &= \frac{0.4 \times 0.6}{50} \left(\frac{200-50}{200-1} \right) \\ &= 0.0048 \times 0.75 = 0.0036 \end{aligned}$$

$$\sigma_{\bar{p}} = \sqrt{\sigma_p^2} = 0.06$$

- أحسب احتمال أن تكون نسبة العينة ($\bar{p}$) أصغر أو تساوي 42%.

$$\begin{aligned} p(\bar{p} \leq 0.42) &= p\left(Z \leq \frac{\bar{p} - \mu_{\bar{p}}}{\sigma_{\bar{p}}}\right) \\ &= p\left(Z \leq \frac{0.42 - 0.40}{0.06}\right) \\ &= p(z \leq 0.33) = 0.6293 \end{aligned}$$

- أحسب احتمال أن تكون نسبة العينة ($\bar{p}$) محصورة ما بين 43% و 45%.

$$\begin{aligned} p(0.43 \leq \bar{p} \leq 0.45) &= p\left(\frac{0.43-0.40}{0.06} \leq Z \leq \frac{0.45-0.40}{0.06}\right) \\ &= p(0.5 \leq Z \leq 0.83) \\ &= \Phi(0.83) - \Phi(0.5) \\ &= 0.7967 - 0.6914 = 0.1053 \end{aligned}$$

مصطلحات الدرس باللغة الانجليزية

Random sample	العينة العشوائية	Sampling distribution	توزيع المعاينة
Population parameters	معالم المجتمع	The Sample Mean	الوسط الحسابي للعينة
The Sample Standard deviation	الانحراف المعياري للعينة	sampling	المعاينة
Sampling distribution of means	توزيع المعاينة للمتوسطات	Population distribution	مجتمع طبيعي
Infinite	مجتمع لا نهائي	Normal distribution	توزيع طبيعي

التقييم الذاتي للدرس الثاني: توزيع المعاينة للمتوسط وتوزيع المعاينة للنسبة

أولاً: أسئلة للفهم

- السؤال 1 : ما المقصود بـ "توزيع المعاينة للمتوسط الحسابي"؟ ولماذا تختلف متوسطات العينات عن بعضها البعض؟
- السؤال 2 : ما الفرق بين السحب بالإرجاع والسحب بدون إرجاع من حيث تأثير كل منهما على تباين توزيع المعاينة؟
- السؤال 3 : ما الفرق بين توزيع ستيودنت (t) والتوزيع الطبيعي المعياري (Z)؟ ومتى نستخدم كلاً منهما؟
- السؤال 4 : ما هي شروط التقريب الطبيعي لتوزيع المعاينة للنسبة؟ ولماذا هذه الشروط ضرورية؟
- السؤال 5 : كيف يتغير الانحراف المعياري لتوزيع المعاينة ($\sigma_{\bar{x}}$) عندما:

- تضاعف حجم العينة ($n \rightarrow 2n$)
- تضاعف تباين المجتمع ($2\sigma^2 \rightarrow 2\sigma^2$)

ثانياً: أسئلة تحليلية

- السؤال 1 : باحث يريد دراسة متوسط دخل الأسر في مدينة جزائرية، قام بسحب عينة عشوائية حجمها $n=100$ وحصل على $\bar{X} = 45000$ دينار، إذا علمت أن الانحراف المعياري للمجتمع $\sigma = 8000$ دينار، فما هو الخطأ المعياري لمتوسط العينة؟ ماذا يعني هذا الرقم؟ وهل يمكن الوثوق بهذا التقدير؟ ولماذا؟
- السؤال 2 : لديك مجتمع حجمه $N=500$ ، وسحبت منه عينة حجمها $n=50$ بدون إرجاع.
- احسب معامل الاستقصاء (n/N) ماذا تستنتج؟
 - هل تحتاج إلى استخدام معامل تصحيح المجتمع المحدود؟ ولماذا؟
- السؤال 3 : في أي من الحالتين التالية يكون التقريب الطبيعي لتوزيع المعاينة للنسبة أكثر دقة؟ ولماذا؟
- الحالة الأولى $n=100, p=0.5$
 - الحالة الثانية $n=100, p=0.05$
- السؤال 4 : ما الفرق بين الحالتين التاليتين من حيث استخدام التوزيع المناسب لحساب الاحتمالات؟
- الحالة A مجتمع طبيعي، σ معلوم، $n=15$
 - الحالة B مجتمع طبيعي، σ مجهول، $n=15$

ثالثاً: تطبيق**التمرين التطبيقي 1:**

المعطيات : مجتمع يتبع التوزيع الطبيعي، متوسطه $\mu = 50$ ، وانحرافه المعياري $\sigma = 10$ تم سحب عينة عشوائية حجمها $n = 25$.

المطلوب:

1. احسب متوسط توزيع المعاينة $(\mu_{\bar{x}})$ والخطأ المعياري $(\sigma_{\bar{x}})$.
2. ما هو احتمال أن يكون متوسط العينة أقل من 47؟
3. ما هو احتمال أن يكون متوسط العينة أكبر من 53؟
4. ما هو احتمال أن يقع متوسط العينة بين 48 و 52؟

التمرين التطبيقي 2:

المعطيات : مجتمع محدود حجمه $N = 1000$ ، متوسطه $\mu = 100$ ، وانحرافه المعياري $\sigma = 15$ تم سحب عينة عشوائية حجمها $n = 100$ بدون إرجاع.

المطلوب:

1. احسب تباين توزيع المعاينة $(\sigma_{\bar{x}}^2)$ والخطأ المعياري $(\sigma_{\bar{x}})$ مع استخدام معامل التصحيح.
2. قارن النتيجة مع الحالة التي نهمل فيها معامل التصحيح، ماذا تلاحظ؟

التمرين التطبيقي 3:

المعطيات : مجتمع يتبع التوزيع الطبيعي، متوسطه المفروض $\mu = 30$ سحبت عينة عشوائية حجمها $(n=16)$ وكانت قيمها كالتالي:

29, 31, 28, 32, 30, 33, 27, 31, 29, 32, 28, 30, 31, 29, 32, 30

المطلوب:

1. احسب متوسط العينة $(\bar{x})$
2. احسب الانحراف المعياري للعينة (S)
3. احسب قيمة t لهذه العينة إذا أردنا اختبار الفرضية أن $\mu = 30$
4. باستخدام جدول توزيع t ، أوجد الاحتمال $P(T > |t|)$ بدرجات حرية مناسبة.

التمرين التطبيقي 4:

المعطيات : أظهرت دراسة سابقة أن 40% من سكان مدينة يستخدمون الهاتف الذكي في التسوق عبر الإنترنت $(p = 0.4)$ تم سحب عينة عشوائية حجمها $n = 200$ شخص.

المطلوب:

1. احسب متوسط توزيع المعاينة للنسبة $(\mu_{\bar{p}})$ والخطأ المعياري $(\sigma_{\bar{p}})$

2. تحقق من شروط التقريب الطبيعي. هل يمكن استخدام التقريب الطبيعي؟ ولماذا؟
3. ما هو احتمال أن تكون نسبة مستخدمي الهاتف الذكي في العينة بين 35% و 45%؟
4. ما هو احتمال أن تتجاوز النسبة 50%؟

التمرين التطبيقي 5:

المعطيات: مصنع أنتج $N = 5000$ قطعة، وكانت نسبة القطع المعيبة في الإنتاج الكلي $p = 0.05$ اشترى تاجر عينة حجمها $n = 400$ قطعة بدون إرجاع.

المطلوب:

1. احسب تباين توزيع المعاينة للنسبة $(\sigma_{\bar{p}}^2)$ والخطأ المعياري $(\sigma_{\bar{p}})$
2. ما هو احتمال أن تكون نسبة القطع المعيبة في العينة أقل من 0.04؟
3. ما هو احتمال أن تكون نسبة القطع المعيبة في العينة بين 0.045 و 0.055؟

رابعاً: تقييم ذاتي سريع

ضع دائرة حول الإجابة الصحيحة:

❖ إذا كان المجتمع يتبع التوزيع الطبيعي بـ $\mu=50$ و $\sigma=10$ ، وحجم العينة $n=25$ ، فإن $\sigma_{\bar{x}}$ يساوي:

A. 10

B. 5

C. 2

D. 0.4

❖ قاعدة أن $n \geq 30$ كافية لتقريب توزيع المعاينة إلى التوزيع الطبيعي ترتبط بـ:

A. قانون الأعداد الكبيرة

B. نظرية النهاية المركزية (CLT)

C. توزيع ستودنت

D. معامل تصحيح المجتمع المحدود

❖ عند حساب احتمالات توزيع المعاينة للمتوسط وكان تباين المجتمع مجهولاً، نستخدم:

A. التوزيع الطبيعي المعياري (Z)

B. توزيع ستودنت (t)

C. توزيع كاي مربع (χ^2)

D. توزيع فيشر (F)

❖ إذا كان متوسط توزيع المعاينة للمتوسط $\mu_{\bar{X}} = \mu$ ، فهذا يعني أن $\bar{X}$ هو مقدر:

- A. متحيز (Biased)
- B. كاف (Sufficient)
- C. غير متحيز (Unbiased)
- D. متنسق (Consistent) فقط

الفصل الاول نظرية المعاينة وتوزيعاتها

الدرس الثالث: توزيع المعاينة للفروق والتباين

المقدمة التعريفية:

في هذا الدرس الثالث، سننتقل من دراسة عينة واحدة إلى مقارنة عينتين أو أكثر، وكذلك إلى دراسة توزيع تباين العينة، سُميَ بين توزيع المعاينة للفروق بين متوسطين وتوزيع المعاينة للفروق بين نسبتيين، المتوسط والتباين والانحراف المعياري للفروق بين متوسطي عينتين مستقلتين، سُنطبق نظرية النهاية المركزية على توزيع الفروق لحساب الاحتمالات، ونُحلل خصائص توزيعات مربع كاي (χ^2) و F و t المرتبطة بتباين العينة، وأخيراً سنُفان بين توزيعات المعاينة المختلفة (الفروق بين متوسطين، الفروق بين نسبتيين، النسبة بين تباينين)، هذا التدرج يمكنكم من التعامل مع مسائل المقارنة بين مجتمعين ومسائل التباين، وهي من أكثر التطبيقات شيوعاً في البحث العلمي وتحليل البيانات.

الهدف العام للدرس

أن يُتقن الطالب بنهاية هذا الدرس تطبيق مفاهيم توزيع المعاينة للفروق (بين متوسطين ونسبتيين) وتوزيع المعاينة للتباين (χ^2 , F , t) على مجتمعات حقيقية أو افتراضية، وذلك من خلال حساب معالم هذه التوزيعات، وحساب الاحتمالات المرتبطة بها، والتمييز بين الحالات التي يستخدم فيها كل توزيع.

الأهداف التفصيلية للدرس

- حساب الطالب المتوسط الحسابي، التباين، والانحراف المعياري للفروق بين متوسطي عينتين مستقلتين، وكذلك للفروق بين نسبتي عينتين مستقلتين؛
- ❖ يُطبق الطالب التوزيع الطبيعي المعياري (Z) لحساب احتمالات الفروق بين متوسطي عينتين ونسبتي عينتين؛
- ❖ يُحدد الطالب درجات الحرية وقيمة إحصائية مربع كاي (χ^2) لتباين عينة واحدة، ويُفسر الاحتمالات المرتبطة بها من جدول التوزيع. ويُحدد درجات حرية البسط والمقام، ويُقدر الاحتمال المقابل من جدول توزيع F .

مواضيع الدرس

- ❖ توزيع المعاينة للفروق بين متوسطين؛
- ❖ توزيع المعاينة للفروق بين نسبتيين؛
- ❖ توزيع المعاينة للتباين (χ^2)؛
- ❖ توزيع النسبة بين تباينين (F -Distribution).

الوقت المتوقع: 4.5 ساعات بين 3 ساعات محاضرات و 1 ساعة و 30 دقيقة أعمال موجهة

تمهيد:

في الدرسين السابقين، تعلمنا كيف يتوزع متوسط العينة ($\bar{X}$) ونسبة العينة ($\bar{p}$) عند سحب عينة واحدة من مجتمع واحد، وكيف يمكننا استخدام هذه التوزيعات لتقدير معالم المجتمع واختبار الفرضيات حول عينة واحدة، لكن في كثير من التطبيقات العملية، لا نكتفي بدراسة مجتمع واحد، بل نحتاج إلى مقارنة مجتمعين أو أكثر، على سبيل المثال: هل متوسط دخل الذكور يختلف عن متوسط دخل الإناث؟ هل نسبة النجاح في مدرسة ما تختلف عن نسبة النجاح في مدرسة أخرى؟ هل تباين إنتاج مصنع يساوي تباين إنتاج مصنع منافس؟

هذه الأسئلة تقودنا إلى دراسة توزيع المعاينة للفروق بين متوسطين و توزيع المعاينة للفروق بين نسبتيين . كما أننا سنتوسع في دراسة توزيع المعاينة للتباين نفسه، والذي يختلف تمامًا عن توزيع المتوسط، حيث يتبع توزيعات خاصة مثل مربع كاي (χ^2) و توزيع F ، هذه التوزيعات تشكل الأدوات الأساسية لاختبار الفرضيات المتعلقة بالتباين ومقارنة تباينين، وهي ضرورية في مجالات مثل مراقبة الجودة، التحليل المالي، والتجارب الطبية.

1- توزيع المعاينة للفرق بين متوسطين

إذا سحبنا عينة 1 بحجم n_1 من المجتمع 1 الذي متوسطه الحسابي μ_1 وانحرافه المعياري σ_1 ، وكان $\bar{X}_1$ متغير توزيع المعاينة للمتوسطات 1.

وسحبنا عينة 2 بحجم n_2 من المجتمع 2 الذي متوسطه الحسابي μ_2 وانحرافه المعياري σ_2 ، وكان $\bar{X}_2$ متغير توزيع المعاينة للمتوسطات 2.

مع العلم أن المجتمعين مستقلين، فإن: $(\bar{X}_1 - \bar{X}_2)$ يمثل متغيراً عشوائياً للفرق بين متوسطي العينتين 1 و 2 ، (بن سبع، 2020، صفحة 33) حيث:

المتوسط الحسابي للفرق $(\bar{X}_1 - \bar{X}_2)$

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2$$

التباين للفرق $(\bar{X}_1 - \bar{X}_2)$

$$\sigma_{\bar{X}_1 - \bar{X}_2}^2 = \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2$$

$$\sigma_{\bar{X}_1 - \bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

الانحراف المعياري للفرق $(\bar{X}_1 - \bar{X}_2)$

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\sigma_{\bar{X}_1 - \bar{X}_2}^2}$$

$$= \sqrt{\sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2}$$

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

مثال تطبيقي 1:

لدينا مجتمعان مستقلان X_1 حيث: متوسطه الحسابي $\mu_1 = 30$ وتباينه $\sigma_1^2 = 25$ المجتمع X_2 الذي متوسطه الحسابي $\mu_2 = 20$ وتباينه $\sigma_2^2 = 16$. قمنا بسحب عينتان حجمهما على التوالي: 30 و 35.

المطلوب: أحسب المتوسط الحسابي، التباين والانحراف المعياري للفرق $(\bar{X}_1 - \bar{X}_2)$ حيث: $\bar{X}_1$ و $\bar{X}_2$ متغيري توزيع المعاينة للمتوسطات 1 و 2 على الترتيب.

الحل:

المتوسط الحسابي للفرق $(\bar{X}_1 - \bar{X}_2)$

$$\begin{aligned}\mu_{\bar{X}_1 - \bar{X}_2} &= \mu_{\bar{X}_1} - \mu_{\bar{X}_2} \\ &= \mu_1 - \mu_2 \\ &= 30 - 20 = 10\end{aligned}$$

التباين للفرق $(\bar{X}_1 - \bar{X}_2)$

$$\begin{aligned}\sigma_{\bar{X}_1 - \bar{X}_2}^2 &= \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \\ &= \frac{25}{30} + \frac{16}{35} = 1.29\end{aligned}$$

الانحراف المعياري للفرق $(\bar{X}_1 - \bar{X}_2)$

$$\begin{aligned}\sigma_{\bar{X}_1 - \bar{X}_2} &= \sqrt{\sigma_{\bar{X}_1 - \bar{X}_2}^2} \\ &= \sqrt{1.29} = 1.136\end{aligned}$$

نكتب:

$$(\bar{X}_1 - \bar{X}_2) \sim N(\mu_{\bar{X}_1 - \bar{X}_2}, \sigma_{\bar{X}_1 - \bar{X}_2})$$

$$(\bar{X}_1 - \bar{X}_2) \sim N(10, 1.136) \text{ نكتب:}$$

طبيعة توزيع المعاينة للفروق والمجاميع:

1* إذا كان المجتمعين يتبعان التوزيعا طبيعيا فان توزيع المعاينة للفروق بين متوسطي عينتين يتبع أيضا التوزيع الطبيعي.

2* إذا كان حجم العينتين (n_1, n_2) أكبر من 30 توزيع المعاينة للفروق بين متوسطي عينتين يتبع التوزيع الطبيعي. (نظرية النهاية المركزية) (تمار ، 2020 ، صفحة 39).
نكتب:

$$(\bar{X}_1 - \bar{X}_2) \sim N(\mu_{\bar{X}_1 - \bar{X}_2}, \sigma_{\bar{X}_1 - \bar{X}_2})$$

أو:

$$(\bar{X}_1 - \bar{X}_2) \sim N(\mu_{\bar{X}_1} - \mu_{\bar{X}_2}, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}})$$

أو كذلك :

$$(\bar{X}_1 - \bar{X}_2) \sim N(\mu_1 - \mu_2, \frac{\sigma_1}{\sqrt{n_1}} + \frac{\sigma_2}{\sqrt{n_2}})$$

حساب الاحتمال في توزيع المعاينة للفروق والمجاميع:

$$Z < \frac{(\bar{x}_1 - \bar{x}_2) - \mu_{\bar{X}_1 - \bar{X}_2}}{\sigma_{\bar{X}_1 - \bar{X}_2}}$$

$$Z < \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

من نفس المثال السابق:

احسب الاحتمال:

$$P((\bar{X}_1 - \bar{X}_2) < 12) = p\left(Z < \frac{(\bar{x}_1 - \bar{x}_2) - \mu_{\bar{X}_1 - \bar{X}_2}}{\sigma_{\bar{X}_1 - \bar{X}_2}}\right)$$

$$= p\left(Z < \frac{(12) - 10}{1.136}\right) = \Phi(1.76) = 0.9608$$

مثال تطبيقي 2:

لمعرفة الفرق بين متوسطي الأجور للعمال، سُحبت عينتين عشوائيتين من مؤسستين مختلفتين لإنتاج مواد التجميل، فتحصلنا على النتائج:

المؤسسة B	المؤسسة A	
180	230	المتوسط الحسابي
1600	1296	التباين
50	36	حجم العينة المسحوبة

- أوجد احتمال أن يكون الفرق بين متوسطي الأجور للعينتين أكبر من 40
الحل:

الطريقة الأولى:

- نبحث أولاً عن معالم توزيع المعاينة للفرق

المتوسط الحسابي للفرق $(\bar{X}_1 - \bar{X}_2)$

$$\begin{aligned}\mu_{\bar{X}_1 - \bar{X}_2} &= \mu_{\bar{X}_1} - \mu_{\bar{X}_2} \\ &= \mu_1 - \mu_2 \\ &= 230 - 180 = 50\end{aligned}$$

التباين للفرق $(\bar{X}_1 - \bar{X}_2)$

$$\begin{aligned}\sigma_{\bar{X}_1 - \bar{X}_2}^2 &= \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \\ &= \frac{1296}{36} + \frac{1600}{50} = 68\end{aligned}$$

الانحراف المعياري للفرق $(\bar{X}_1 - \bar{X}_2)$

$$\begin{aligned}\sigma_{\bar{X}_1 - \bar{X}_2} &= \sqrt{\sigma_{\bar{X}_1 - \bar{X}_2}^2} \\ &= \sqrt{68} = 8.24\end{aligned}$$

$$(\bar{X}_1 - \bar{X}_2) \sim N(\mu_{\bar{X}_1 - \bar{X}_2}, \sigma_{\bar{X}_1 - \bar{X}_2})$$

نكتب:

$$(\bar{X}_1 - \bar{X}_2) \sim N(50, 8.24)$$

- ثم نحسب الاحتمال في المرحلة الثانية:

$$\begin{aligned}
p((\bar{X}_1 - \bar{X}_2) > 40) &= p\left(Z > \frac{(\bar{x}_1 - \bar{x}_2) - \mu_{\bar{X}_1 - \bar{X}_2}}{\sigma_{\bar{X}_1 - \bar{X}_2}}\right) \\
&= p\left(Z > \frac{(40) - 50}{8.24}\right) \\
&= p(Z > -1.21) \\
&= \Phi(1.21) = 0.8869
\end{aligned}$$

الطريقة الثانية: نحسب الاحتمال باستخدام المتغير المعياري Z مباشرة

$$\begin{aligned}
p((\bar{X}_1 - \bar{X}_2) > 40) &= p\left(Z > \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right) \\
&= p\left(Z > \frac{40 - (230 - 180)}{\sqrt{\frac{1296}{36} + \frac{1600}{50}}}\right) \\
&= p\left(Z > \frac{40 - (230 - 180)}{\sqrt{\frac{1296}{36} + \frac{1600}{50}}}\right) \\
&= p\left(Z > \frac{-10}{\sqrt{68}}\right) \\
&= p\left(Z > \frac{-10}{\sqrt{68}}\right) \\
&= p(Z > -1.21) \\
&= p(Z < 1.21) = 0.8869
\end{aligned}$$

مثال تطبيقي 3:

متوسط مدة حياة المصابيح التي تنتجها مؤسسة جزائرية هو 325 يوم ($\mu_1 = 325$)، بانحراف معياري ($\sigma_1 = 9$)، أما متوسط مدة حياة المصابيح التي تنتجها مؤسسة صينية هو 322 يوم ($\mu_2 = 322$)، بانحراف معياري ($\sigma_2 = 7$).

تم أخذ عيّنتين من المصابيح التي تنتجها المؤسستين الجزائرية والصينية حجم كل منهما ($n_1 = n_2 =$) (36).

المطلوب:

- حساب متوسط توزيع المعاينة للفرق بين متوسطي العينتين.
- حساب تباين توزيع المعاينة للفرق بين متوسطي العينتين.
- حساب الانحراف المعياري لتوزيع المعاينة للفرق بين متوسطي العينتين.
- حساب الاحتمالات التالية:

$$p((\bar{X}_1 - \bar{X}_2) \leq 10) ; \quad p((\bar{X}_1 - \bar{X}_2) \geq 0) ; \quad p(5 \leq (\bar{X}_1 - \bar{X}_2) \leq 8)$$

الحل: لدينا:

$$\mu_1 = 325$$

$$\mu_2 = 322$$

$$\sigma_1 = 9$$

$$\sigma_2 = 7$$

$$n_1 = n_2 = 36$$

بما أن $n_1, n_2 \geq 30$

فان توزيع المعاينة للفرق بين متوسطي العينتين $((\bar{X}_1 - \bar{X}_2))$ يتبع التوزيع الطبيعي.

- حساب متوسط توزيع المعاينة للفرق بين متوسطي العينتين.

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_{\bar{X}_1} - \mu_{\bar{X}_2}$$

$$= \mu_1 - \mu_2$$

$$= 325 - 322 = 3$$

- حساب تباين توزيع المعاينة للفرق بين متوسطي العينتين.

$$\sigma_{\bar{X}_1 - \bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

$$= \frac{9^2}{36} + \frac{7^2}{36} = 3.61$$

- حساب الانحراف المعياري لتوزيع المعاينة للفرق بين متوسطي العينتين.

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\sigma_{\bar{X}_1 - \bar{X}_2}^2}$$

$$= \sqrt{3.61} = 1.9$$

- حساب الاحتمالات التالية:

$$\begin{aligned}
p((\bar{X}_1 - \bar{X}_2) \leq 10) &= p\left(Z \leq \frac{(\bar{x}_1 - \bar{x}_2) - \mu_{\bar{X}_1 - \bar{X}_2}}{\sigma_{\bar{X}_1 - \bar{X}_2}}\right) \\
&= p\left(Z \leq \frac{(10) - 3}{1.9}\right) \\
&= p(Z \leq 3.68) \\
&= \Phi(3.68) = 0.9998 \\
p((\bar{X}_1 - \bar{X}_2) \geq 0) &= p\left(Z \geq \frac{(0) - 3}{1.9}\right) \\
&= p(Z \geq -1.57) \\
&= p(Z \leq +1.57) \\
&= \Phi(1.57) = 0.9417 \\
&= p\left(\frac{5-3}{1.9} \leq Z \leq \frac{(8)-3}{1.9}\right) = p(1.05 \leq Z \leq 2.63) = \Phi(2.63) - \Phi(1.05) \\
&= 0.9957 - 0.8531 = 0.1426
\end{aligned}$$

مثال تطبيقي 3:

إذا كان متوسط الدخل الشهري للعائلات بالمدينة (أ) هو 16500، مع انحراف معياري قدره: 2500 دج، ومتوسط الدخل الشهري للعائلات بالمدينة (ب) هو 15750 دج، مع انحراف معياري قدره: 2000 دج. وسحبنا عينتين عشوائيتين من المدينتين (أ) و (ب) حجمهما على التوالي: 125 عائلة، و 100 عائلة.

1. أوجد احتمال أن يكون متوسط الدخل للعينة الأولى أكبر من متوسط الدخل للعينة الثانية بأكثر من 330 دج.

2. أوجد احتمال أن يكون متوسط الدخل للعينة الأولى أقل من متوسط الدخل للعينة الثانية.

$$\begin{aligned}
p((\bar{X}_1 - \bar{X}_2) > 330) &= p\left(Z > \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right) \\
&= 0.9192 \\
\Phi(1.4) &= 0.9192 \quad \Phi(2.5) = 0.9938
\end{aligned}$$

$$p((\bar{X}_1 - \bar{X}_2) < 0) = p\left(Z < \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right) = 0.0062$$

2-توزيع المعاينة للفرق بين نسبتين

ليكن لدينا مجتمعان يتبعان التوزيع ذو الحدين، (توزيع برنولي).

إذا سحبنا عينة 1 بحجم n_1 من المجتمع 1 الذي نسبة النجاح فيه p_1 وانحرافه المعياري σ_{p_1} ، وكان $\bar{p}_1$ متغير توزيع المعاينة لمتوسطات النسب 1.

وسحبنا عينة 2 بحجم n_2 من المجتمع 2 الذي نسبة النجاح فيه p_2 وانحرافه المعياري σ_{p_2} ، وكان $\bar{p}_2$ متغير توزيع المعاينة لمتوسطات النسب 2.

مع العلم أن المجتمعين مستقلين، فإن: $(\bar{p}_1 - \bar{p}_2)$ يمثل متغيرا عشوائيا للفرق بين متوسطي العينتين 1 و 2 ويقترّب من التوزيع الطبيعي، حيث (Prem & Christopher, 2010, p. 329):

المتوسط الحسابي للفرق في النسب $(\bar{p}_1 - \bar{p}_2)$

$$\mu_{\bar{p}_1 - \bar{p}_2} = \mu_{\bar{p}_1} - \mu_{\bar{p}_2} = p_1 - p_2$$

التباين للفرق في النسب $(\bar{p}_1 - \bar{p}_2)$

$$\sigma_{\bar{p}_1 - \bar{p}_2}^2 = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$$

$$\sigma_{\bar{p}_1 - \bar{p}_2}^2 = \sigma_{\bar{p}_1}^2 + \sigma_{\bar{p}_2}^2$$

الانحراف المعياري للفرق في النسب $(\bar{p}_1 - \bar{p}_2)$

$$\sigma_{\bar{p}_1 - \bar{p}_2} = \sqrt{\sigma_{\bar{p}_1 - \bar{p}_2}^2} = \sqrt{\sigma_{\bar{p}_1}^2 + \sigma_{\bar{p}_2}^2}$$

مثال تطبيقي 1:

إذا كانت نسبة النجاح في مدرسة للبنات هي 0.7، وكانت نسبة النجاح في مدرسة للذكور هي 0.65، وأخذنا عينة عشوائية من المدرستين بحجم 70 و 35 على التوالي، فما احتمال أن لاتزيد نسبة النجاح في الفرق بين العينتين عن 0.1؟

الحل:

بما أن المجتمعين يتبعان توزيع برنولي (لأن الظاهرة المدروسة وهي نسبة النجاح تقبل نتيجتين نجاح / فشل)

وبالتالي:

نرمز لنسبة النجاح في مدرسة البنات $p_1 = 0.7$

نرمز لنسبة عدم النجاح في مدرسة البنات $q_1 = 1 - p_1 = 0.3$

نرمز لنسبة النجاح في مدرسة الذكور $p_2 = 0.35$

نرمز لنسبة عدم النجاح في مدرسة الذكور $q_2 = 1 - p_2 = 0.65$

المتوسط الحسابي للفرق في النسب $(\bar{p}_1 - \bar{p}_2)$

$$\begin{aligned}\mu_{\bar{p}_1 - \bar{p}_2} &= p_1 - p_2 \\ &= 0.7 - 0.35 = 0.35\end{aligned}$$

التباين للفرق في النسب $(\bar{p}_1 - \bar{p}_2)$

$$\begin{aligned}\sigma_{\bar{p}_1 - \bar{p}_2}^2 &= \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2} \\ &= \frac{0.7 * 0.3}{70} + \frac{0.35 * 0.65}{35} \\ &= 0.003 + 0.0065 = 0.0095\end{aligned}$$

الانحراف المعياري للفرق في النسب $(\bar{p}_1 - \bar{p}_2)$

$$\sigma_{\bar{p}_1 - \bar{p}_2} = \sqrt{0.0095} = 0.097$$

لحساب الاحتمال نستخدم المتغير المعياري القياسي Z

$$\left(Z = \frac{(\bar{p}_1 - \bar{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \right)$$

• ثم نحسب الاحتمال في المرحلة الثانية:

$$\begin{aligned}p((\bar{p}_1 - \bar{p}_2) \leq 0.1) &= p\left(Z \leq \frac{(\bar{p}_1 - \bar{p}_2) - \mu_{\bar{p}_1 - \bar{p}_2}}{\sigma_{\bar{p}_1 - \bar{p}_2}}\right) \\ &= p\left(Z \leq \frac{(0.1) - 0.35}{0.097}\right) \\ &= p(Z \leq -2.57) \\ &= 1 - \Phi(2.57) \\ &= 1 - 0.9949 = 0.0051\end{aligned}$$

الطريقة الثانية: نحسب الاحتمال باستخدام المتغير المعياري z مباشرة

$$\begin{aligned}
p((\bar{p}_1 - \bar{p}_2) \leq 0.1) &= p\left(Z = \frac{(\bar{p}_1 - \bar{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}\right) \\
&= p\left(Z \leq \frac{0.1 - (0.7 - 0.35)}{\sqrt{\frac{0.7*0.3}{70} + \frac{0.35*0.65}{35}}}\right) \\
&= p\left(Z \leq \frac{-0.25}{\sqrt{0.003 + 0.0065}}\right) \\
&= p\left(Z \leq \frac{-0.25}{\sqrt{0.097}}\right) \\
&= p(Z \leq -2.57) \\
&= 1 - p(Z \leq 2.57) \\
&= 1 - \Phi(2.57) \\
&= 1 - 0.9949 = 0.0051
\end{aligned}$$

مثال تطبيقي 2:

- مجتمع مكون من 70 شخص من بينهم 42 شخص مصاب بضغط الدم، أوجد:
- التوزيع العيني لنسبة المصابين بضغط الدم في عينة عشوائية حجمها 10 مسحوبة من هذا المجتمع.
 - ثم احسب احتمال أن تزيد نسبة المصابين في العينة 43%.
- الحل:

$$p = 0.6 \quad \text{alors} \quad q = 0.4$$

بما أن $n > 30$ فإن توزيع المعاينة يقترب من التوزيع الطبيعي حيث:

$$\bar{p} \sim N(\mu_{\bar{p}}, \sigma_{\bar{p}})$$

الوسط الحسابي لتوزيع معاينة النسبة

$$\mu_{\bar{p}} = p = 0.6$$

التباين لتوزيع معاينة النسبة

السحب بدون إرجاع

$$n/N = 0.14 > 0.05$$

نستخدم معامل التصحيح

$$\begin{aligned}\sigma_{\bar{p}}^2 &= \frac{pq}{n} \left(\frac{N-n}{N-1} \right) \\ &= \frac{0.6 \times 0.4}{10} \left(\frac{70-10}{70-1} \right) \\ &= 0.0208\end{aligned}$$

الانحراف المعياري لتوزيع معاينة النسبة

$$\begin{aligned}\sigma_{\bar{p}} &= \sqrt{\frac{pq}{n} \left(\frac{N-n}{N-1} \right)} \\ &= \sqrt{0.0208} = 0.1442 \\ \bar{p} &\sim N(0.6, 0.1442)\end{aligned}$$

- حساب الاحتمال:

$$\begin{aligned}p(\bar{p} > 0.43) &= p\left(Z > \frac{\bar{p} - \mu_{\bar{p}}}{\sigma_{\bar{p}}}\right) \\ &= p\left(Z > \frac{0.43 - 0.6}{0.1442}\right) \\ &= p(z > -1.17) \\ &= p(z < 1.17) = 0.8790\end{aligned}$$

نسبة الناجحين في كليتي الاقتصاد والحقوق على التوالي هي : 35% و 40% على التوالي، نسحب عينتين من الكليتين على الترتيب بحجم 40 و 50 المطلوب:

- أحسب احتمال $p((\bar{p}_1 - \bar{p}_2) \leq 0.05)$

الحل:

$$\begin{aligned}p((\bar{p}_1 - \bar{p}_2) \leq 0.05) &= p\left(Z = \frac{(\bar{p}_1 - \bar{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}\right) \\ &= p\left(Z \leq \frac{0.05 - (0.35 - 0.40)}{\sqrt{\frac{0.35 \times 0.65}{40} + \frac{0.40 \times 0.60}{50}}}\right)\end{aligned}$$

$$\begin{aligned}
&= p\left(Z \leq \frac{0.1}{\sqrt{0.0056 + 0.0048}}\right) \\
&= p\left(Z \leq \frac{0.1}{0.1019}\right) \\
&= p(Z \leq 0.98) = 0.8365
\end{aligned}$$

3-توزيع المعاينة للتباين

يُعتبر تباين العينة S^2 من أهم الإحصاءات في الإحصاء الاستدلالي، حيث يُستخدم كمقدر لتباين المجتمع σ^2 وعلى عكس متوسط العينة ($\bar{x}$) الذي يقترب توزيعه المعاني من التوزيع الطبيعي بموجب نظرية النهاية المركزية، فإن لتوزيع المعاينة لتباين العينة خصائص مختلفة تمامًا، سنتناول في هذا العنصر من الدرس الأسس النظرية لتوزيع المعاينة للتباين، بما في ذلك خصائصه للمجتمعات الطبيعية، وتوزيع مربع كاي (χ^2)، وتوزيع F، وتوزيع t حيث تشكل هذه التوزيعات العمود الفقري للعديد من إجراءات اختبار الفرضيات وبناء فترات الثقة في الإحصاء التطبيقي (Casella & Berger, 2024).

الحالة العامة: المجتمعات غير الطبيعية

ليكن S^2 هو تباين عينة عشوائية حجمها n مأخوذة من مجتمع متوسطه μ وتباينه σ^2 ، وعزمه الرابع حول المتوسط هو μ_4 حيث:

$$\mu_{S^2} = \sigma^2 \quad \text{and} \quad \sigma^2_{S^2} = \frac{\mu_4 - \sigma^2}{n-1}$$

وفقاً لـ (Montgomery & Runger, 2019, p. 245)، فإن تباين توزيع المعاينة لـ S^2 يُعطى بالعلاقة:

$$Var(S^2) = \frac{\mu_4}{n} - \frac{n-3}{n(n-1)} \sigma^4$$

حيث:

- μ_4 هو العزم الرابع حول المتوسط
- σ^4 هو مربع تباين المجتمع
- n هو حجم العينة

الحالة الخاصة: المجتمع الطبيعي

عندما يتبع المجتمع التوزيع الطبيعي، فإن العزم الرابع حول المتوسط يُعطى بالعلاقة:

$\mu_4 = 3\sigma^4$ (Hogg, Tanis, & Zimmerman, 2024, p. 189) وبتعويض هذه القيمة في المعادلة

السابقة، نحصل على:

$$\text{Var}(S^2) = \frac{3\sigma^4}{n} - \frac{n-3}{n(n-1)}\sigma^4 = \frac{2\sigma^4}{n-1}$$

ملاحظة هامة : S^2 لا يتوزع توزيعاً طبيعياً حتى لو كان المجتمع الطبيعي الأصلي، ولكنه يقترب من التوزيع الطبيعي عندما يكون حجم العينة n كبيراً .

توزيع مربع كاي (Chi-Square Distribution)

جاءت فكرة توزيع كاي تربيع من نظرية التوزيع الطبيعي المعياري، حيث إذا كانت $(X_1, X_2, X_3, \dots, X_n)$ متغيرات عشوائية مستقلة لكل منها توزيع طبيعي معياري، فإن مجموع مربعات هذه المتغيرات يرمز له بالرمز: $\chi^2 = X_1^2 + X_2^2 + X_3^2 \dots + X_n^2$ ، وتوزيع هذا المجموع هو ما يعرف بتوزيع "كاي تربيع" وبالتالي عندما نقول أن المتغير (X) يتبع توزيع "كاي تربيع" فهذا يعني أن هذا المتغير هو عبارة عن مجموع مربعات (n) من المتغيرات العشوائية المستقلة والتي يخضع كل منها للتوزيع الطبيعي، أي أن عدد درجات الحرية (n) يتبع عدد المتغيرات، ولهذا العدد (n) أهمية كبيرة تظهر في أثناء استخدام جداول (χ_n^2) ، (ديب، 2009، صفحة 209) وهذه الجداول تعطي قيم الاحتمالات التالية:

$$P(X = x) = \int_{x_0}^{+\infty} f(x)dx$$

تعريف توزيع مربع كاي

يعتبر توزيع مربع كاي من التوزيعات الأساسية في الإحصاء التطبيقي. دالة كثافته الاحتمالية (PDF) تُعطى بالعلاقة التالية :

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$$

باستخدام هذه الصيغة، يمكن حساب دالة التوزيع التراكمية لتوزيع كاي تربيع لقيم مختلفة من x و n . هناك جداول خاصة بهذا التوزيع تبين قيم χ_n^2 النظرية، (القيم الحرجة)، عند درجة حرية معينة ومستوى معنوية يرمز له بـ α (أنظر الملحق رقم 02)، فمثلاً لإيجاد قيم χ_n^2 عند درجة حرية 7 ومستوى معنوية $\alpha = 0.005$ نبحث عنها في الجدول، وهي نقطة التقاء السطر 7 مع العمود $\chi_{0.005}^2$ ، نجدها تساوي 20.278.

الشكل رقم 1: كيفية البحث عن قيم χ_n^2 في جدول توزيع كاي تربيع.

df	$\chi_{0.100}^2$	$\chi_{0.050}^2$	$\chi_{0.025}^2$	$\chi_{0.010}^2$	$\chi_{0.005}^2$
1	2.706	3.841	5.024	6.635	7.879
2	4.605	5.991	7.378	9.210	10.597
3	6.251	7.815	9.348	11.345	12.838
4	7.779	9.488	11.143	13.277	14.860
5	9.236	11.070	12.833	15.086	16.750
6	10.645	12.592	14.449	16.812	18.548
7	12.017	14.067	16.013	18.475	20.278
8	13.362	15.507	17.535	20.090	21.955

المصدر: من إعداد الباحث.

تمثل القيمة 20.278 قيمة المتغير العشوائي χ_n^2 التي تعطي احتمال متراكم قدره 0.995 للمجال $[0 - 20.278]$ ، أو هي قيمة المتغير العشوائي χ_n^2 التي تعطي احتمال متراكم قدره 0.005 للمجال $[20.278 - (+\infty)]$. وهذا يعني أن:

$$P(X = x) = \int_0^{\chi_n^2} f(x^2) dx^2 = 1 - \alpha$$

أو:

$$P(X = x) = \int_{\chi_n^2}^{+\infty} f(x^2) dx^2 = \alpha$$

أما إن كان المجتمع الأصلي يخضع للتوزيع الطبيعي فإن المتغير $\frac{(n-1)S^2}{\sigma^2}$ يخضع لتوزيع يسمى

توزيع مربع كاي χ^2 بعدد درجات حرية يساوي $(n - 1)$. أي أن

$$\chi^2 = \frac{(n - 1)S^2}{\sigma^2} \sim \chi_{n-1}^2$$

ودالة كثافته تعطى بالشكل التالي:

$$f(y) = y^{\frac{v-1}{2}} e^{-y/2}, \quad y > 0$$

حيث v هي درجات الحرية، وتكون القيمة المتوقعة لهذا التوزيع هي v و تباينه هو $2v$ أي بمعنى أن:

$$E(X) = \mu = n \text{ التوقع الرياضي}$$

إثبات هذه العلاقة انطلاقاً من المتغير (X) الذي يتبع توزيع "كاي تربيع" بدرجة حرية (n) .
نضع: $X = \sum_{i=1}^n z_i^2$ حيث (z_i) متغير عشوائي طبيعي معياري.
من خصائص التوزيع الطبيعي المعياري:

$$E(z) = 0$$

$$\sigma_z^2 = 1$$

من تعريف القيمة المتوقعة:

$$E(X) = \int_0^1 x f(x) dx$$

$$\begin{aligned} E(X) &= E(\sum_{i=1}^n z_i^2) \\ &= \sum_{i=1}^n E(z_i^2) \\ &= \sum_{i=1}^n \sigma_z^2 \\ &= \sum_{i=1}^n 1 \\ &= n \end{aligned}$$

التباين والانحراف المعياري:

التباين يعطى بالعلاقة التالية: $\sigma^2 = 2n$

أما الانحراف المعياري: $\sigma = \sqrt{2n}$

أثبت أن: $\sigma^2 = 2n$

$$\sigma^2 = \sigma_{\sum_{i=1}^n z_i^2}^2$$

نعلم أن: $\sigma_{z^2}^2 = 2$

$$\begin{aligned} &= \sum \sigma_{z^2}^2 \\ &= \sum 2 \\ &= 2n \end{aligned}$$

4-توزيع النسبة بين تباينين

إذا كان (S_1^2) هو تباين عينه عشوائية حجمها n_1 مسحوبة من مجتمع طبيعي $N(\mu_1, \sigma_1)$ ، وكان (S_2^2) هو تباين عينه عشوائية أخرى حجمها n_2 ومسحوبة من مجتمع طبيعي آخر $N(\mu_2, \sigma_2)$ وكانت العينتان مستقلتان فإن المتغير (LeBlanc, 2004, p. 236):

$$\frac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$$

حيث أن $F(n_1 - 1, n_2 - 1)$ تسمى بتوزيع F بدرجة الحرية $n_1 - 1$ و $n_2 - 1$ و دالة الكثافة الاحتمالية للمتغير y الذي يخضع لتوزيع F بدرجة الحرية v_1, v_2 :

$$f(y) = \frac{y^{\frac{v_1}{2}-1}}{(v_1 y + v_2)^{\frac{v_1+v_2}{2}}}, \quad y > 0$$

مثال تطبيقي 1:

تم سحب عينة عشوائية حجمها $n=10$ من مجتمع يتبع التوزيع الطبيعي، ووجد أن تباين العينة هو $S^2=15$ المطلوب:

- احسب قيمة إحصائية مربع كاي (χ^2) إذا علمت أن تباين المجتمع الحقيقي هو $\sigma^2=10$.
- ما هي درجات حرية هذا التوزيع؟
- أوجد احتمال أن يكون تباين العينة أقل من 15 (أي $P(S^2 < 15)$)

الحل:

- حساب قيمة χ^2

نستخدم العلاقة:

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$

$$\chi^2 = \frac{(10-1)15}{10}$$

$$\chi^2 = 13.5$$

- درجات الحرية:

$$V = n - 1$$

$$V = 9$$

- حساب الاحتمال:

$$\begin{aligned} P(S^2 < 15) &= p\left(\frac{(n-1)S^2}{\sigma^2}\right) \\ &= p\left(\frac{(n-1)15}{\sigma^2}\right) \\ &= P(\chi_9^2 < 13.5) \end{aligned}$$

من جدول توزيع مربع كاي بدرجات حرية 9:

- القيمة الحرجة عند مستوى 0.10 هي 14.684

- القيمة الحرجة عند مستوى 0.20 هي 12.242

بما أن 13.5 تقع بين 12.242 و 14.684، فإن:

$$0.10 < P(\chi_9^2 < 13.5) < 0.20$$

وبالتقريب (باستخدام الاستيفاء الخطي أو الآلة الحاسبة):

$$P(S^2 < 15) = P(\chi_9^2 < 13.5) = 0.157$$

مثال تطبيقي 2:

تم سحب عينتين مستقلتين من مجتمعين طبيعيين:

العينة الأولى: حجمها $n_1 = 8$ ، وتباينها $S_1^2 = 12$

العينة الثانية: حجمها $n_2 = 10$ ، وتباينها $S_2^2 = 8$

المطلوب:

- احسب قيمة إحصائية F

- ما هي درجات حرية البسط والمقام؟

- أوجد احتمال أن تكون النسبة بين التباينين $\frac{S_1^2}{S_2^2}$ أقل من القيمة المحسوبة لـ F.

الحل:

حساب قيمة F

$$\frac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$$

$$F = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}}$$

$$= \frac{S_1^2}{S_2^2}$$

$$= \frac{12}{8} = 1.5$$

درجات الحرية:

البسط:

$$v_1 = n_1 - 1 = 8 - 1 = 7$$

المقام:

$$v_2 = n_2 - 1 = 10 - 1 = 9$$

حساب الاحتمال:

$$P(F < 1.5)$$

$$P(F(7,9) < 1.5)$$

من جدول توزيع F بدرجتي حرية (7,9):

القيمة الحرجة عند $\alpha = 0.25$ هي 1.42

القيمة الحرجة عند $\alpha = 0.20$ هي 1.70

بما أن 1.5 تقع بين 1.42 و 1.70، فإن:

$$0.20 < P(F(7,9) < 1.5) < 0.25$$

وبالتقريب:

$$P(F(7,9) < 1.5) \approx 0.22$$

التقييم الذاتي للدرس الثالث: توزيع المعاينة للفروق والتباين

أولاً: أسئلة للفهم

السؤال 1 : عرف المتغير العشوائي $((\bar{x}_1 - \bar{x}_2))$ وماذا يمثل؟ وما هي شروط استقلالية المجتمعين؟

السؤال 2 : اكتب الصيغ الرياضية لكل من:

- متوسط الفرق بين متوسطي عينتين $(\bar{x}_1 - \bar{x}_2)$ - تباين الفرق بين متوسطي عينتين $\sigma_{\bar{x}_1 - \bar{x}_2}^2$

السؤال 3 : متى يمكن القول بأن توزيع المعاينة للفرق بين متوسطي عينتين يتبع التوزيع الطبيعي؟

السؤال 4 : ما هو توزيع مربع كاي (Chi-Square) ؟ وما هي العلاقة التي تربطه بتباين العينة S^2 عندما

يكون المجتمع طبيعياً؟

ثانياً: أسئلة تحليلية

السؤال 1 : لدينا مجتمعان مستقلان طبيعيان:

المجتمع الأول $\mu_1=50$ ، $\sigma_1=10$ المجتمع الثاني $\mu=45$ ، $\sigma_1=8$ سحبت عينتان حجمهما $n_1=25$ ، $n_2=36$ هل توزيع $(\bar{x}_1 - \bar{x}_2)$ طبيعي؟ ولماذا؟ احسب متوسطه وتباينه.السؤال 2 : لماذا لا يتوزع تباين العينة S^2 توزيعاً طبيعياً حتى لو كان المجتمع طبيعياً؟ وما التوزيع

الذي يتبعه بدقة؟

السؤال 3 : في اختبار مقارنة تباينين، إذا كانت قيمة F المحسوبة قريبة من 1، ماذا يمكنك أن تستنتجحول العلاقة بين σ_1^2 و σ_2^2 ؟ وإذا كانت F كبيرة جداً (أكبر من 3) ماذا يعني ذلك؟

ثالثاً: تطبيق

التمرين التطبيقي 1:

لدينا مجتمعان مستقلان X_1 و X_2 حيث:

$$X_2 \sim N(12, 14)$$

$$X_1 \sim N(10, 12)$$

ما هو التوزيع العني للفرق بين متوسطي عينتين، مسحوبتان من X_2 و X_1 حجمهما على التوالي: 36 و 45؟

ما هو احتمال أن لا يزيد الفرق عن 5؟

التمرين التطبيقي 2:

البيانات التالية متعلقة بمعدلات طالبة السنة الثانية في الكليتين A و B

الكلية B	الكلية A	
13.25	11.5	المتوسط الحسابي
4	2.25	الانحراف المعياري
74	42	حجم العينة المسحوبة

المطلوب:

حساب الاحتمال:

$$p(\bar{X}_1 - \bar{X}_2) \leq -2.5)$$

التمرين التطبيقي 3:

نسبة الرياضيين في مدينة ما هي 30%، نسحب عينة عشوائية بحجم 140 شخص من سكان هذه المدينة.

المطلوب:

- أحسب احتمال أن لا تزيد نسبة الرياضيين في العينة عن 20%.
- أحسب احتمال أن لا تقل نسبة الرياضيين في العينة عن 70%.
- أحسب احتمال أن لا تقل نسبة غير الرياضيين في العينة عن 40%.
- أحسب احتمال أن تتراوح نسبة الرياضيين في العينة ما بين 25% و 30%.

التمرين التطبيقي 4:

نسبة الناجحين في كليتي الاقتصاد والحقوق على التوالي هي : 75% و 80% على التوالي، نسحب عينتين من الكليتين على الترتيب بحجم 80 و 150

المطلوب:

- أحسب احتمال $p((\bar{p}_1 - \bar{p}_2) \leq 0.05)$
- أحسب احتمال $p((\bar{p}_1 - \bar{p}_2) > 0.55)$
- أحسب احتمال $p((\bar{p}_1 - \bar{p}_2) \leq 0.25)$
- أحسب احتمال $p((\bar{p}_1 - \bar{p}_2) \leq 0.75)$
- أحسب احتمال $p(0.15 \leq (\bar{p}_1 - \bar{p}_2) \leq 0.25)$

التمرين التطبيقي 5:

تم سحب عینتين مستقلتين من مجتمعين طبيعيين:

العينة الأولى: حجمها $n_1 = 18$ ، وتباينها $S_1^2 = 3.5$

العينة الثانية: حجمها $n_2 = 22$ ، وتباينها $S_2^2 = 2.7$

المطلوب:

- احسب قيمة إحصائية F
- ما هي درجات حرية البسط والمقام؟
- أوجد احتمال أن تكون النسبة بين التباينين $\frac{S_1^2}{S_2^2}$ أقل من القيمة المحسوبة لـ F.

الفصل الثاني نظرية التقدير

الدرس الرابع: نظرية التقدير وطرق إيجاد المقدرات (طريقة العزوم والمعقولية العظمى) المقدمة التعريفية:

في هذا الدرس ننقل من دراسة توزيعات المعاينة إلى جوهر الإحصاء الاستدلالي: كيف نستخدم بيانات العينة لتقدير معالم المجتمع المجهولة؟ سنعرّف المفاهيم الأساسية مثل المعلمة والمقدر والتقدير، ثم نطبق طريقتين منهجيتين لاشتقاق المقدرات هما طريقة العزوم وطريقة المعقولية العظمى.

الهدف العام للدرس

أن يُتقن الطالب بنهاية هذا الدرس تطبيق طريقتي العزوم والمعقولية العظمى لاشتقاق مقدرات المعالم المجهولة في مجتمعات إحصائية بسيطة، وذلك من خلال التمييز بين المفاهيم الأساسية، وحساب المقدرات خطوة بخطوة، وتقييم مزايا وعيوب كل طريقة، والتحقق من شروط الاستخدام.

الأهداف التفصيلية للدرس

- ❖ يُعرّف الطالب كلاً من: المعلمة (Parameter)، المقدر (Estimator)، التقدير (Estimate)؛
- ❖ يُطبق الطالب خطوات طريقة العزوم (Method of Moments) لإيجاد مقدرات معالم المجتمع؛
- ❖ يُحلل الطالب دالة الاحتمال ولو غاريتها كخطوة أساسية في طريقة المعقولية العظمى، ويستخدم التفاضل لاشتقاق مقدر المعقولية العظمى (MLE) لمتوسط مجتمع طبيعي (بتباين معلوم)؛

مواضيع الدرس

- ❖ المفاهيم الأساسية في نظرية التقدير؛
- ❖ طريقة العزوم (Method of Moments)؛
- ❖ طريقة المعقولية العظمى (Maximum Likelihood)؛
- ❖ تطبيق عملي على طريقة MLE.

الوقت المتوقع: 4.5 ساعات بين 3 ساعات محاضرات و 1 سا و 30 حصة أعمال موجهة

تمهيد

في الدروس السابقة، درسنا توزيعات المعاينة المختلفة (للمتوسط، النسبة، التباين، الفروق) وكيفية حساب الاحتمالات المرتبطة بها، لكن السؤال الجوهرى في الإحصاء الاستدلالي هو: كيف نستخدم بيانات العينة لتقدير معالم المجتمع (مثل μ ، σ^2 ، p)؟ هذا هو موضوع نظرية التقدير (Estimation Theory). سنقتصر في هذا الدرس على المفاهيم الأساسية فقط، ثم ننتقل إلى طريقتين أساسيتين لاشتقاق المقدرات (Estimators):

– طريقة العزوم (Method of Moments) ؛

– طريقة المعقولة العظمى (Maximum Likelihood Method).

1- نظرية التقدير: المفاهيم الأساسية

كما نعلم أن المعلمة (Parameter) هي قيمة عددية ثابتة (غالبًا مجهولة) تصف المجتمع الإحصائي، مثل: متوسط المجتمع (μ)، تباين المجتمع (σ^2)، نسبة النجاح في المجتمع (p)، معامل ارتباط المجتمع (ρ)، نرسم للمعلمة عادة بحروف إغريقية (مثل θ , μ , σ , p). (عاروري، 2013، صفحة 75).

المقدر (Estimator)

المقدر (Estimator) هو قاعدة أو دالة رياضية تُطبق على بيانات العينة للحصول على تقدير عددي للمعلمة المجهولة. يُرمز له عادة بـ $\hat{\theta}$. مثال :

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \text{ هو مقدر للمعلمة } \mu$$

التقدير (Estimate)

عندما نطبق المقدر على بيانات عينة فعلية، نحصل على قيمة عددية تسمى التقدير (Estimate). مثال: إذا كانت العينة: {5, 7, 9}، فإن :

$$\bar{X} = \frac{\sum x_i}{N} = 7 \text{ هو تقدير لـ } \mu.$$

الجدول رقم 2: الفرق بين المقدر والتقدير.

المقدر (Estimator)	التقدير (Estimate)
متغير عشوائي (قبل سحب العينة)	قيمة عددية ثابتة (بعد سحب العينة)
له توزيع معاينة وتباين	ليس له توزيع (رقم واحد)
مثال: $\bar{X}$	مثال: $\bar{X} = \frac{\sum x_i}{N} = 7$

المصدر: من اعداد الباحث.

عند مواجهة معلمة مجهولة (مثل μ أو σ^2 أو p أو أي معلمة أخرى في توزيع غير مألوف)، نحتاج إلى طريقة منهجية لبناء مقدر مناسب من العينة، هناك عدة طرق، أبرزها:

- طريقة العزوم (Method of Moments) ؛
- طريقة المعقولية العظمى (Maximum Likelihood) ؛
- طريقة المربعات الصغرى - (Least Squares) للإحصاء الخطي؛
- طريقة بيز (Bayes Method) .

سنركز على طريقتي العزوم والمعقولية العظمى.

2- طريقة العزوم (Method of Moments)

تعتمد طريقة العزوم على مبدأ بسيط: مساواة عزوم المجتمع (النظرية) بعزوم العينة (التجريبية)، ثم حل المعادلات للحصول على مقدرات المعالم. (GIRI, Narayan, 1993, p. 285)

العزم حول الصفر من الرتبة k للمجتمع (النظري):

$$\mu'_k = E[X^k]$$

العزم حول الصفر من الرتبة k للعينة (التجريبية):

$$m'_k = \frac{1}{n} \sum_{i=1}^n X_i^k$$

خطوات طريقة العزوم

نفترض أن لدينا معالم مجهولة عددها r مثل $(\theta_1, \theta_2, \dots, \theta_r)$ ، نكتب أول عزوم نظرية للمجتمع بدلالة هذه المعالم

$$\mu'_1 = g_1(\theta_1, \theta_2, \theta_3, \dots, \theta_r), \mu'_2 = g_2(\theta_1, \theta_2, \theta_3, \dots, \theta_r), \dots$$

نساوي كل عزم نظري بالعزم التجريبي المقابل من العينة:

$$\mu'_1 = m'_1, \quad \mu'_2 = m'_2, \quad \dots$$

نحل نظام المعادلات الناتج للحصول على مقدرات المعالم $((\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \dots, \hat{\theta}_r))$ (Mátyás, 1999, p. 4)

مميزات وعيوب طريقة العزوم

ميزة: بسيطة وسهلة الحساب، ودائماً تعطي حلاً (غالباً).

عيب: ليست دائماً الأكثر كفاءة (أكبر تباين) مقارنة بطريقة المعقولية العظمى.

3- طريقة المعقولية العظمى (Maximum Likelihood)

دالة الاحتمال (Likelihood Function)

إذا كانت دالة الاحتمال (Likelihood Function)

إذا كانت $X_1, X_2, \dots, X_n$ عينة عشوائية من مجتمع له دالة كتلة احتمال (للمنفصل) أو دالة كثافة (للمتصل) تعتمد على معلمة θ ، فإن دالة الاحتمال (Likelihood) تُعرف بأنها دالة المعلمة θ الناشئة عن العينة:

$$L(\theta) = f(x_1; \theta) \cdot f(x_2; \theta) \dots f(x_n; \theta)$$

$$= \prod_{i=1}^n f(x_i; \theta)$$

حيث: العينة ثابتة، $X_1, X_2, \dots, X_n$ متغير، و θ ثابت.

نبحث عن قيمة θ التي تُعظم دالة الاحتمال $L(\theta)$ ؛ أي القيمة التي تجعل البيانات المرصودة أكثر احتمالاً (أكثر "معقولة").

مقدر المعقولة العظمى (MLE – Maximum Likelihood Estimator)

$$\hat{\theta}_{MLE} = \arg \max_{\theta \in \Theta} L(\theta)$$

أي قيمة θ التي ترفع $L(\theta)$ إلى أقصى حد ممكن.

خطوات طريقة MLE للتوزيعات المتصلة (ذات دالة كثافة) (GIRI, Narayan, 1993, p. 287)

- كتابة دالة الاحتمال:

$$L(\theta) = \prod_{i=1}^n f(x_i; \theta)$$

- أخذ اللوغاريتم الطبيعي (لتبسيط الضرب إلى جمع، خاصة مع التوزيع الأسّي:

$$\ell(\theta) = \ln L(\theta) = \sum_{i=1}^n \ln f(x_i; \theta)$$

- إيجاد مشتقة $\ell(\theta)$ بالنسبة لـ θ أو مشتقات جزئية إذا كان θ متجهًا ومساواتها بالصفر:

$$\frac{\partial \ell(\theta)}{\partial(\theta)} = 0$$

- حل المعادلة للحصول على $\hat{\theta}_{MLE}$

- للتحقق التأكد أن المشتقة الثانية سالبة (عظمى) (مصطفى، 2002، الصفحات 158-180).

مثال تطبيقي:

في المجتمع الطبيعي بوسط حسابي (μ) وتباين معلوم ((σ^2))، حيث $X_i \sim N(\mu, \sigma^2)$

دالة الكثافة الاحتمالية:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

المطلوب:

- أوجد المقدّر المبني على التوقع الأعظمي للوسط (μ) (باستخدام طريقة الإمكان الأكبر).

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \dots\dots\dots (1)$$

وبإدخال اللوغاريتم على طرفي المعادلة (1) نحصل على:

$$\boxed{0.5} \quad \log f(x) = \log\left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\right)$$

$$\log f(x) = \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{1}{2\sigma^2}(x - \mu)^2$$

نأخذ المجموع:

$$\sum_{i=1}^n \log f(x) = \sum_{i=1}^n \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x - \mu)^2$$

$$\boxed{0.5} \quad \sum_{i=1}^n \log f(x) = n \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x - \mu)^2$$

وباشتقاق طرفي المعادلة أعلاه بالنسبة ل μ نحصل على:

$$\frac{\partial}{\partial \mu} \sum_{i=1}^n \log f(x) = 0 + \frac{1}{\sigma^2} \sum_{i=1}^n (x - \mu)$$

وعند مساواة المعادلة للصفر نحصل على:

$$\frac{\partial}{\partial \mu} \sum_{i=1}^n \log f(x) = 0 \Rightarrow \frac{1}{\sigma^2} \sum_{i=1}^n (x - \mu) = 0$$

$$\frac{1}{\sigma^2} \neq 0 \rightarrow \sum_{i=1}^n (x - \mu) = 0$$

$$\sum_{i=1}^n x_i - n\mu = 0 \Rightarrow \boxed{\mu = \frac{\sum_{i=1}^n x_i}{n}}$$

المقدر المبني على التوقع الأعظمي للوسط (μ) هو:

$$\mu = \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

مميزات وعيوب طريقة MLE

المميزات:

- تتمتع بخصائص مقاربة ممتازة (للعينات الكبيرة): اتساق، كفاءة مقاربة، طبيعية مقاربة؛
- عادة تنتج مقدرات ذات تباين أصغر من طريقة العزوم (للعينات الكبيرة)؛
- مفهومها بديهي: نختار المعلمة التي تجعل البيانات المرصودة أكثر احتمالاً.

العيوب:

- قد تؤدي إلى معادلات معقدة لا حل تحليلي لها (تحتاج إلى حل عددي)؛
- تعتمد على افتراض معرفة صيغة التوزيع الاحتمالي بالكامل؛
- يمكن أن تكون حساسة للقيم المتطرفة (Outliers) .

التقييم الذاتي للدرس الرابع: نظرية التقدير – طريقة العزوم والمعقولية العظمى

أولاً: أسئلة للفهم

- ما الفرق بين "المعلمة" و "المقدر"؟
- متى نسمي القيمة العددية الناتجة من تطبيق المقدر على عينة فعلية بـ "التقدير"؟
- ما المبدأ الأساسي الذي تقوم عليه طريقة العزوم؟
- ما هي دالة الاحتمال (Likelihood) في طريقة المعقولية العظمى؟
- لماذا نأخذ اللوغاريتم لدالة الاحتمال قبل الاشتقاق؟

ثانياً: أسئلة تحليلية

- كيف تتصرف طريقة العزوم إذا كان لدينا معلمتان مجهولتان (مثل المتوسط والتباين معاً)؟
- لماذا تُعتبر طريقة المعقولية العظمى أكثر كفاءة من طريقة العزوم في العينات الكبيرة؟
- ما المشكلة التي قد تواجهها طريقة MLE إذا كان شكل التوزيع معقداً جداً؟
- هل تعتمد طريقة العزوم على معرفة الشكل الدقيق لتوزيع المجتمع؟ ولماذا؟

ثالثاً: تطبيق بسيط

التمرين 1:

لديك مجتمع يتبع توزيع برنولي (نتيجتان: نجاح/فشل). أردت تقدير نسبة النجاح في المجتمع (p) باستخدام طريقة العزوم:

- ما هو العزم النظري الأول لهذا التوزيع؟
- ما هو العزم التجريبي الأول من العينة؟
- كيف نحصل على مقدر p حسب طريقة العزوم؟

التمرين 2:

لديك نفس المجتمع (برنولي)، تريد الآن استخدام طريقة المعقولية العظمى لتقدير p ما هي دالة الاحتمال التي نريد تعظيمها؟
بأي خطوة بسيطة يُمكن الحصول على مقدر p؟
رابعاً: تقييم ذاتي سريع (اختر الإجابة الصحيحة)

المعلمة (Parameter) هي:

أ) دالة في بيانات العينة

ب) قيمة ثابتة (غالباً مجهولة) تصف المجتمع

ج) نتيجة تطبيق مقدر على عينة

- (د) نوع من المتغيرات العشوائية
- مقدر المعقولية العظمى (MLE) يتم إيجاده عن طريق:**
- (أ) مساواة العزوم النظرية بالتجريبية
- (ب) تعظيم دالة الاحتمال
- (ج) إيجاد أقل مجموع مربعات
- (د) استخدام معاملات بايز
- طريقة العزوم تقوم على:**
- (أ) مساواة العزوم النظرية بالعزوم التجريبية
- (ب) تعظيم دالة اللوغاريتم
- (ج) تحليل التباين
- (د) اختبار المعنوية
- لإيجاد مقدر بطريقة MLE ، غالبًا نأخذ اللوغاريتم لدالة الاحتمال لأن:**
- (أ) يحول الضرب إلى جمع فيسهل الاشتقاق
- (ب) لأن النتائج تكون دائمًا أكبر
- (ج) لأن اللوغاريتم يجعل الدالة خطية دائمًا
- (د) لأن دالة الاحتمال تكون دائمًا سالبة
- أي مما يلي يُعد ميزة لطريقة المعقولية العظمى؟**
- (أ) لا تحتاج إلى معرفة شكل التوزيع
- (ب) دائمًا بسيطة جدًا وتُعطي حلًا مغلقًا
- (ج) تتمتع بخصائص مقاربة ممتازة (للعينة الكبيرة)
- (د) غير حساسة للقيم المتطرفة
- مثال على مقدر تم اشتقاقه بطريقة العزوم في الدرس لمجتمع برنولي هو:**
- (أ) متوسط العينة
- (ب) نسبة النجاح في العينة
- (ج) انحراف معياري العينة
- (د) الفرق بين متوسطين
- التقدير (Estimate) هو:**
- (أ) متغير عشوائي
- (ب) قاعدة رياضية عامة

- (ج) قيمة عددية من عينة فعلية
- (د) معلمة المجتمع الحقيقية
- عند اشتقاق مقدار المتوسط (μ) لمجتمع طبيعي باستخدام MLE ، كانت النتيجة:
- (أ) متوسط العينة
- (ب) انحراف معياري العينة
- (ج) تباين العينة
- (د) قيمة ثابتة تساوي صفرًا

الفصل الثاني نظرية التقدير

الدرس الخامس: خصائص المقدرات النقطية (في العينات الصغيرة والعينات الكبيرة)

المقدمة التعريفية:

في هذا الدرس الخامس، سننتقل من مرحلة إيجاد المقدرات إلى مرحلة تقييمها واختيار أفضلها. سنعرّف التقدير النقطي ونميزه عن تقدير الفترة، سنحدد خمس خصائص أساسية للمقدرات الجيدة: عدم التحيز، الاتساق، الكفاءة، الكفاية، والتقارب، ونحلل سلوك هذه الخصائص في سياق العينات الصغيرة مقابل العينات الكبيرة، سنقارن بين مقدرين أو أكثر لنفس المعلمة باستخدام معيار الكفاءة (التباين)، سنقيم أي المقدرات أكثر ملاءمة وفقاً لحجم العينة المتاح.

الهدف العام للدرس

أن يتقن الطالب بنهاية هذا الدرس تقييم المقدرات النقطية وفقاً لخصائصها الرئيسية (عدم التحيز، الاتساق، الكفاءة، الكفاية، والتقارب) والتمييز بين سلوك هذه الخصائص في العينات الصغيرة والكبيرة، وذلك من خلال التحقق من تحقق الشروط الرياضية لكل خاصية، والمقارنة بين مقدرات مختلفة، واختيار المقدر الأنسب بناءً على حجم العينة.

الأهداف التفصيلية للدرس

- ❖ يُعرّف الطالب التقدير النقطي؛
- ❖ يستنتج الطالب شروط تحقق خاصية عدم التحيز (Unbiasedness) وخاصية الاتساق (Consistency) لمقدر معين؛
- ❖ يُحلل الطالب مفهوم الكفاية (Sufficiency) ويُفسر معناه؛
- ❖ يُقارن الطالب بين سلوك المقدرات (خاصة التقارب Convergence) في العينات الصغيرة والعينات الكبيرة من حيث دقة التقدير وسرعة الاقتراب من المعلمة الحقيقية.

مواضيع الدرس

- ❖ أنواع التقدير؛
- ❖ تعريف التقدير النقطي (Point Estimation)؛
- ❖ خصائص المقدرات الجيدة .

الوقت المتوقع: 4.5 ساعات بين 3 ساعات محاضرات و 1 سا و 30 حصة أعمال موجهة

تمهيد:

نظرية التقدير في الإحصاء هي فرع من فروع الإحصاء التي تهتم بتطوير طرق وأساليب لتقدير معلومة أو خاصية مجموعة من البيانات الإحصائية الكبيرة بناءً على عينة محدودة من هذه المجموعة. هدف نظرية التقدير هو إيجاد طريقة تقديرية موثوقة ودقيقة للكشف عن خصائص البيانات الكبيرة دون الحاجة إلى تحليل أو عملية إحصائية للمجموعة بأكملها. تعتمد نظرية التقدير على استخدام نماذج وأدوات إحصائية لتحديد العلاقة بين العينة والمجموعة الكبيرة ومن ثم توفير تقديرات للخصائص المطلوبة. يستخدم هذا النوع من التحليل في العديد من التطبيقات الإحصائية مثل استطلاعات الرأي والدراسات العينية وتحليل السكان وغيرها.

1- التقدير النقطي

التقدير النقطي هو أسلوب إحصائي يُستخدم لتقدير إحدى معالم المجتمع (مثل المتوسط، التباين، أو النسبة) بقيمة واحدة (نقطة واحدة) تمثل أفضل تخمين لتلك المعلمة استناداً إلى بيانات العينة. لنفترض أنك تريد تقدير متوسط عمر المصابيح الكهربائية التي ينتجها مصنع معين، كيف يمكنك فعل ذلك؟ (ابو صالح، 2020، صفحة 255).

- تقوم بسحب عينة عشوائية من إنتاج المصنع.
 - تجري اختبارات على المصابيح المختارة لتسجيل عمر كل مصباح حتى احتراقه.
 - تحسب المتوسط الحسابي لأعمار المصابيح في العينة.
- هذا المتوسط المحسوب من العينة هو **تقدير نقطي** لمتوسط عمر المصابيح في المجتمع بأكمله (كل إنتاج المصنع). وببساطة، نستخدم قيمة واحدة (وسط العينة) كتقدير لمعلمة المجتمع (وسط المجتمع). كما سبق أن درسنا في موضوع التوزيعات المعيارية، فإن متوسط العينة ($\bar{X}$) يُعد مقدراً غير متحيز لمتوسط المجتمع (μ)، وهو أحد الأمثلة الشائعة على التقدير النقطي.

أنواع التقدير في الإحصاء:

يتم استخدام التقدير النقطي والتقدير بمجال الثقة لتقدير القيمة المجهولة أو المعلمة في عينة معينة (الهوبي، 2014، صفحة 62).

أ. التقدير النقطي (Point estimation):

يُستخدم التقدير النقطي لتمثيل معلمة المجتمع بقيمة ما نحسبها من خلال مقدر لعينة عشوائية.

ب. التقدير بمجال الثقة (Interval estimation):

على عكس التقدير النقطي، يُستخدم التقدير بمجال الثقة لتقدير نطاق التغطية المحتمل للمعلمة المجهولة. يُستخدم عادةً مستوى ثقة معين، مثل 95%، لتحديد نطاق القيم المرجحة للمعلمة. على سبيل المثال، يمكن استخدام الانحراف المعياري و عينة الحجم لتحديد نطاق ثقة لمتوسط المجتمع باستخدام توزيع ت ستودنت.

بشكل عام، يتم استخدام التقدير النقطي والتقدير بمجال الثقة فيا لإحصاء لتحديد المعلومات واتخاذ القرارات على أساس العينات المأخوذة من المجتمع الكلي.

2- بعض خصائص المقدّر

عند الرغبة في تقدير معالم مجتمع معين، من الضروري اختيار الإحصائية المناسبة من العينة لتقدير تلك المعالم. تُستخدم الإحصائية المقابلة في العينة عادةً كأفضل تقدير ممكن للمعلمة المستهدفة في المجتمع. فمثلاً، يُقدر متوسط المجتمع (μ) بمتوسط العينة ($\bar{X}$)، وتُسمى الإحصائية المستخدمة في هذه العملية بالمقدر (Estimator).

فيما يلي أبرز الخصائص التي يُفضّل أن تتوفر في المقدّر الإحصائي الجيد (عاروري، 2013، صفحة 76):

المقدر غير المتحيز (Unbiasedness)

يُعرف المقدّر بأنه غير متحيز إذا كانت القيمة المتوقعة له (متوسط تقديراته عبر عينات متعددة) تساوي القيمة الحقيقية للمعلمة في المجتمع.

يمكن التعريف الرياضي لخاصية المقدّر غير متحيز إذا كانت ($\hat{\theta}$) هي المقدّر للمعلمة (θ)، فإن المقدّر غير متحيز إذا كان:

$$E(\hat{\theta}) = \theta$$

كما يُفضل التقدير الذي يتمتع بأقل متوسط مربع خطأ (MSE) بين المقدرات غير المتحيزة (بفضيل كمال، 2025).

الخصائص:

- لا يبالغ المقدّر في تقدير المعلمة أو يقلل من شأنها في المتوسط.
- خطأ التحيز يساوي صفراً، حيث أن الفرق بين متوسط التقديرات والقيمة الحقيقية معدوم.
- تعتبر هذه الخاصية مهمة، لكنها لا تكفي وحدها للحكم على جودة المقدّر (قد يكون غير متحيز وتباينه عالٍ جداً (جميل أحمد، 2010)).

خاصية عدم الاتساق (Consistency):

يقال بأن $\hat{\theta}$ مقدراً متنسقاً لمعلمة المجتمع θ إذا كان $\hat{\theta}$ يميل إلى θ (أي تقترب منها) كلما زاد حجم العينة، ويكون المقدّر متنسقاً إذا كان:

$$\begin{cases} \lim_{n \rightarrow \infty} E(\hat{\theta}) = \theta \\ \lim_{n \rightarrow \infty} V(\hat{\theta}) = 0 \end{cases}$$

بمعنى أن احتمال أن يكون المقدّر قريباً جداً من المعلمة الحقيقية يقترب من الواحد عندما يكبر حجم العينة (بلفضيل كمال، 2025).

خاصية الكفاءة (Efficiency):

الكفاءة هي خاصية تقارن بين مقدرين غير متحيزين أو أكثر، المقدّر الأكثر كفاءة (الأكفاً) هو الذي يمتلك أقل تباين (Minimum Variance) بين جميع المقدرات غير المتحيزة. المقدّر ذو التباين الأقل يعني أن تقديراته تتجمع بشكل أكبر حول القيمة الحقيقية، مما يقلل من خطأ التقدير. قد يوجد للمعلمة الواحدة أكثر من مقدر ويمكننا المقارنة بين هذه المقدرات من خلال المقارنة بين تبايناتهم حيث نعتبر أن المقدّر الأقل تبايناً هو المقدّر الأكثر كفاءة، فإذا كان لدينا المقدران $\hat{\theta}_1$ و $\hat{\theta}_2$ للمعلمة θ ، فإن المقدّر $\hat{\theta}_1$ أكثر كفاءة من المقدّر $\hat{\theta}_2$ إذا تحقق ما يلي:

$$e = \frac{V(\hat{\theta}_1)}{V(\hat{\theta}_2)} < 1$$

خاصية الكفاية (Sufficiency):

يقال أن المقدّر $\hat{\theta}$ مقدراً كافياً للمعلمة θ إذا كان التقدير $\hat{\theta}$ يحتوي على كل المعلومات الموجودة في العينة عن المعلمة θ وهذا يعني أنه بعد معرفة $\hat{\theta}$ فإن المعلومات المتبقية في العينة لا تفيد في معرفة θ . وهي عملية التقدير لجميع المشاهدات الممكنة للعينة أو على الأقل أغلبها (الهوبي، 2014، الصفحات 63-65).

خاصية التقارب:

خاصية التقارب (Convergence) هي من أهم الخصائص المرغوبة في المقدرات الإحصائية. تعني أن قيمة المقدّر تقترب من قيمة المعلمة الحقيقية للمجتمع كلما زاد حجم العينة n بمعنى آخر، كلما كبرت العينة، أصبح التقدير أكثر دقة وأقل انحرافاً عن الحقيقة.

نقول أن المقدّر $\hat{\theta}_1$ أكثر تقارباً من $\hat{\theta}_2$ في تقديره للمعلمة θ إذا كان احتمال أو درجة الثقة للأولى أكبر من الثانية أي:

$$P(\theta - \lambda \langle \hat{\theta}_1 \leq \theta + \lambda \rangle) > P(\theta - \lambda \langle \hat{\theta}_2 \leq \theta + \lambda \rangle)$$

للتوضيح يمكن افتراض أنك تريد تقدير متوسط طول طلاب جامعة، تأخذ عينة صغيرة (20 طالباً) ويختلف متوسطها قليلاً عن الحقيقة، تأخذ عينة أكبر (200 طالب) يصبح متوسطها أقرب، تأخذ عينة ضخمة (2000 طالب) يصبح متوسطها قريباً جداً من المتوسط الحقيقي للمجتمع، هذا هو جوهر خاصية التقارب.

التقييم الذاتي للدرس الخامس: خصائص المقدرات النقطية

أولاً: أسئلة للفهم

- ماذا يُقصد بـ "التقدير النقطي"؟ أعط مثلاً واحداً.
- ما الفرق بين التقدير النقطي والتقدير بفترة (مجال الثقة)؟
- متى نقول عن مقدر ما إنه "غير متحيز"؟
- ماذا تعني خاصية "الاتساق" بالنسبة للمقدر؟
- أي المقدرين تفضل: مقدر تباينه صغير أم مقدر تباينه كبير؟ ولماذا؟

ثانياً: أسئلة تحليلية

- هل وجود خاصية عدم التحيز وحدها كافية لاختيار مقدر جيد؟ لماذا؟
- كيف تتصرف المقدرات المتسقة عندما يزيد حجم العينة إلى أعداد كبيرة جداً؟
- لو كان لديك مقدران غير متحيزين لنفس المعلمة، كيف تقرر أيهما أفضل؟
- ما العلاقة بين خاصية التقارب وحجم العينة؟
- هل يمكن أن يكون مقدر ما متحيزاً لكنه أفضل من مقدر غير متحيز في بعض الحالات؟ كيف؟

ثالثاً: تطبيق لفظي

الموقف 1:

يريد باحث تقدير متوسط دخل الأسر في مدينة، استخدم متوسط عينة مكونة من 30 أسرة.

- هل يمكن اعتبار هذا المقدر "غير متحيز"؟ ولماذا؟

الموقف 2:

باحث آخر استخدم نفس الطريقة لكن على عينة مكونة من 500 أسرة.

- كيف تتغير خاصية "الاتساق" عند مقارنة العينة الصغيرة بالكبيرة؟

الموقف 3:

لديك مقدران لتقدير نسبة المصابين بمرض في مجتمع:

- المقدر الأول: تباينه 0.01

- المقدر الثاني: تباينه 0.04

أيهما أكثر كفاءة؟ ولماذا؟

رابعاً: تقييم ذاتي سريع (اختر العبارة الصحيحة)

التقدير النقطي هو:

(أ) تقدير معلمة المجتمع بقيمة واحدة من العينة

- (ب) تقدير معلمة المجتمع بفترة تحتوي على القيمة الحقيقية
- (ج) تقدير تباين العينة فقط
- (د) تقدير حجم العينة المناسب
- المقدر "غير المتحيز" يعني أن:**
- أ) قيمته دائماً تساوي المعلمة الحقيقية
- (ب) متوسط تقديراته عبر العينات المتكررة يساوي المعلمة الحقيقية
- (ج) تباينه يساوي صفراً
- (د) حجم عينته كبير جداً
- خاصية "الاتساق" تعني أن المقدر:**
- أ) يظل ثابتاً مهما تغيرت العينة
- (ب) يقترب من المعلمة الحقيقية كلما كبر حجم العينة
- (ج) غير متحيز دائماً
- (د) له تباين يساوي صفراً
- المقدر "الأكثر كفاءة" بين مقدرين غير متحيزين هو الذي:**
- أ) له متوسط حسابي أكبر
- (ب) له تباين أصغر
- (ج) يعتمد على عينة أصغر
- (د) له انحراف معياري أكبر
- خاصية التقارب ترتبط ارتباطاً وثيقاً ب:**
- أ) حجم العينة الصغير
- (ب) حجم العينة الكبير
- (ج) شكل توزيع المجتمع فقط
- (د) عدم التحيز فقط

الفصل الثاني نظرية التقدير

الدرس السادس: معلومة فيشر ، Borne de Cramer-Rao

المقدمة التعريفية:

بعد أن تعرفنا في الدرس السابق على خصائص المقدرات الجيدة (عدم التحيز، الاتساق، الكفاءة، الكفاية، التقارب)، يبقى سؤال جوهري: كيف نعرف أن مقدراً معيناً هو الأفضل بين جميع المقدرات غير المتحيزة؟ وما أقل تباين يمكن أن يصل إليه مقدر غير متحيز على الإطلاق؟

في هذا الدرس السادس، سنتعرف على أداة نظرية قوية تجيب عن هذا السؤال، وهي معلومة فيشر (Fisher Information) ومتباينة كرامير - راو (Cramér-Rao Lower Bound).

سنُعرّف معلومة فيشر وأهميتها في قياس كمية المعلومات التي تحملها العينة عن المعلمة المجهولة، سنُحلل متباينة كرامير - راو التي تعطي حداً أدنى لتباين أي مقدر غير متحيز، سنطبق خطوات حساب معلومة فيشر لمشاهدة واحدة ولعينة كاملة، ونستخدمها لتحديد الحد الأدنى للتباين، هذا الدرس يمثل نتيجاً لنظرية التقدير ويؤسس للمفاهيم المتقدمة في الإحصاء الرياضي.

الهدف العام للدرس

أن يُتقن الطالب بنهاية هذا الدرس تطبيق متباينة كرامير - راو لتحديد الحد الأدنى لتباين المقدرات غير المتحيزة، وذلك من خلال حساب معلومة فيشر، ومقارنة تباين المقدرات الفعلية بهذا الحد للحكم على كفاءتها، وتحديد ما إذا كان مقدر معين هو أفضل مقدر غير متحيز (مقدر فعال).

الأهداف التفصيلية للدرس

- ❖ يُعرّف الطالب كلاً من: معلومة فيشر للمشاهدة الواحدة وللعينة، ومتباينة كرامير - راو، والمقدر الفعال (Efficient Estimator).
- ❖ يُطبق الطالب متباينة كرامير - راو لاستنتاج الحد الأدنى لتباين المقدرات غير المتحيزة
- ❖ يُقارن الطالب تباين مقدر شائع (مثل متوسط العينة، نسبة النجاح) بالحد الأدنى لكرامير - راو، ويستنتج ما إذا كان ذلك المقدر فعالاً (يحقق المساواة) أم لا.
- ❖ يُحدد الطالب الحالات التي لا تنطبق فيها شروط المتباينة (مثل التوزيع المنتظم).
- ❖ **مواضيع الدرس**
- ❖ التمهيد والمفاهيم الأساسية؛
- ❖ معلومة فيشر (Fisher Information)؛
- ❖ متباينة كرامير - راو (Cramér-Rao Lower Bound)؛
- ❖ أهمية المتباينة وقيودها.

الوقت المتوقع: 4.5 ساعات بين 3 ساعات محاضرات و 1 ساعة و 30 دقيقة أعمال موجهة

تمهيد:

في نظرية التقدير، بعد أن تعرفنا على خصائص المقدرات الجيدة (عدم التحيز، الاتساق، الكفاءة)، يبرز سؤال مهم: **ما هو أقل تباين يمكن أن يصل إليه مقدر غير متحيز؟** و هل هناك حد أدنى لتباين أي مقدر غير متحيز، وإذا كان موجوداً، فكيف يمكن حسابه؟

هذا السؤال أجابت عنه **متباينة كريمير - راو (Cramér–Rao Lower Bound – CRLB)**، (Nielsen, 2013) التي سميت باسم الإحصائيين السويدي هارالد كريمير والهندي كاليانبادي رادهاكريشنا راو، تعطي هذه المتباينة حدًا أدنى لتباين جميع المقدرات غير المتحيزة، وتعتبر أداة أساسية لتقييم كفاءة المقدرات.

1- المفاهيم الأساسية

قد يكون لدينا مقدران غير متحيزين لنفس المعلمة، المقدر الأفضل هو ذو التباين الأصغر (الأكثر كفاءة)، لكن كيف نعرف أن مقدرًا معينًا قد بلغ أقل تباين ممكن؟ متباينة كريمير - راو تقدم لنا حدًا أدنى نظريًا، ومن خلال مقارنة تباين المقدر بهذا الحد يمكننا الحكم على كفاءته.

حتى تنطبق متباينة كريمير - راو، يجب توفر الشروط التالية (تسمى **شروط الانتظام**):

- المعلمة حقيقية $\theta \in \Theta \subseteq \mathbb{R}$: يمكن تعميمها إلى متجه، لكننا نركز على حالة المعلمة الواحدة.
- دالة الكثافة المشتركة (أو دالة الكتلة) للعينة $f(\mathbf{x}; \theta)$ معرفة ومتصلة.
- مجال الدالة لا يعتمد على θ : أي أن مجموعة قيم X لا تتغير بتغير θ (شرط المجال المستقل).
- قابلية الاشتقاق : يمكن اشتقاق دالة الاحتمال (أو اللوغاريتم) بالنسبة لـ θ تحت علامة التكامل/المجموع.
- التباين محدود : أن يكون تباين المقدر محدودًا، وأن تكون معلومة فيشر (Fisher Information) معرفة وموجودة.

في معظم التوزيعات المعروفة (طبيعي، بواسون، برنولي، أسّي، ...) هذه الشروط متحققة.

2-متباينة كريمير - راو (الصيغة الأساسية) (عبد الحفيظ محمد فوزي و عبد الحفيظ مصطفى، 2000)

إذا كان $\hat{\theta}$ مقدرًا غير متحيز للمعلمة θ أي:

$$E(\hat{\theta}) = \theta$$

وتحقق شروط الانتظام، فإن:

$$Var(\hat{\theta}) \geq \frac{1}{I_n(\theta)}$$

حيث $I_n(\theta)$ هي معلومة فيشر (Fisher Information) للعينة كاملة بحجم (n).

2- معلومة فيشر للعينة

تعرف معلومة فيشر للعينة العشوائية المستقلة $X_1, X_2, \dots, X_n$ بأنها:

$$I_n(\theta) = E \left[\frac{\partial}{\partial \theta} \ln L(X; \theta) \right]^2$$

أي توقع مربع مشتقة لوغاريتم دالة الاحتمال (Likelihood) يمكن أيضاً حسابها بالصيغة البديلة (إذا توفرت قابلية الاشتقاق الثانية):

$$I_n(\theta) = -E \left[\frac{\partial^2}{\partial \theta^2} \ln L(X; \theta) \right]$$

وبفضل الاستقلال، فإن معلومة فيشر للعينة تساوي n مضروباً في معلومة فيشر للملاحظة الواحدة:

$$I_n(\theta) = nI_1(\theta)$$

حيث:

$$I_1(\theta) = E \left[\frac{\partial^2}{\partial \theta^2} \ln L(X; \theta) \right]$$

إذن، تكتب متباينة كريمر – راو غالباً على التالي:

$$Var(\hat{\theta}) \geq \frac{1}{I_n(\theta)}$$

3-المقدر الفعال (Efficient Estimator)

إذا وصل مقدر غير متحيز (θ) إلى الحد الأدنى في متباينة كرامر - راو (أي تحقق المساواة)، فإنه يُسمى **مقدراً فعالاً (Efficient Estimator)** (GIRI, Narayan, 1993, p. 278):

$$Var(\hat{\theta}) \geq \frac{1}{I_n(\theta)}$$

في هذه الحالة، يكون المقدر قد بلغ أقل تباين ممكن بين جميع المقدرات غير المتحيزة، وبالتالي فهو **أفضل مقدر غير متحيز (Best Unbiased Estimator)**.

وماتجدد الإشارة إليه أن ليست كل المعلومات لها مقدر فعال، في بعض التوزيعات، لا يوجد مقدر يصل إلى كرامر - راو، أو قد يكون الحد الأدنى غير قابل للتحقيق.

نسمي الكمية:

$$\frac{1}{n E \left[\frac{\partial}{\partial \theta} \ln L(X; \theta) \right]^2}$$

حد كرامر المتعلق بالوسيط μ .

ينتج من النظرية السابقة أنه إذا كان ($\bar{X}$) مقدراً غير متحيز للوسيط (μ) بحيث تباينه يساوي حد كرامر أي:

$$\sigma_{\bar{X}}^2 = \frac{1}{n E \left[\frac{\partial}{\partial \theta} \ln L(X; \theta) \right]^2}$$

عندئذ لن يوجد أي مقدر آخر غير متحيز لـ (μ) يكون تباينه أصغر من تباين ($\bar{X}$)، وبالتالي وفي هذه الحالة يكون المقدر ($\bar{X}$)، هو أفضل المقدرات غير المتحيزة للوسيط (μ)، وبمعنى آخر، المقدر الذي يصل إلى حد كرامر هو أفضل المقدرات (GIRI, Narayan, 1993, p. 278).

4-أهمية المتباينة في الإحصاء

تقييم الكفاءة : توفر معياراً لمعرفة ما إذا كان مقدر معين هو الأفضل بين المقدرات غير المتحيزة.

تحسين التقدير : إذا لم يصل مقدر إلى الحد الأدنى، يمكن البحث عن مقدر آخر أفضل أو تطبيق طرق مثل تقدير المعقولية العظمى (MLE) الذي غالبًا ما يكون مقاربًا فعالاً.

5- قيود المتباينة

الحالات التي لا تنطبق فيها : مثل المجتمعات التي لا يحقق شرط المجال المستقل (مثال: التوزيع المنتظم $U(0, \theta)$ لأن مجاله يعتمد على θ).

المقدرات المتحيزة : يمكن أن يكون لمقدر متحيز تباين أصغر من حد كريمر – راو (مثال: مقدر انكماشى مثل مقدر جيمس-ستاين)، لكن عندئذٍ يُسمح بالتحيز.

العينات الصغيرة : الحد الأدنى هو نظري لجميع العينات، لكن المقدرات الفعالة قد لا توجد إلا في عائلات معينة (مثل الأسية).

التقييم الذاتي للدرس السادس: معلومة فيشر ومتباينة كرامير – راو

أولاً: أسئلة للفهم

- ماذا تعني "معلومة فيشر"؟ وماذا تقيس؟
- ما هي الفائدة الأساسية من متباينة كرامير – راو؟
- متى نُسَمي مقدرًا غير متحيز بـ "مقدر فعال"؟
- اذكر شرطًا واحدًا من شروط الانتظام التي يجب توفرها لانطباق متباينة كرامير – راو.
- لماذا قد لا تنطبق متباينة كرامير – راو على توزيع منتظم مثل $U(0, \theta)$ ؟

ثانياً: أسئلة تحليلية

- لو كان لديك مقدر غير متحيز وتباينه يساوي تماماً حد كرامير – راو، ماذا تستنتج عنه؟
- هل يمكن لمقدر متحيز أن يكون له تباين أصغر من حد كرامير – راو؟ ولماذا قد نستخدمه رغم التحيز؟

- كيف يساعدنا حد كرامير – راو في مقارنة مقدرين غير متحيزين لنفس المعلمة؟
- ما العلاقة بين حجم العينة وقيمة حد كرامير – راو (هل يزيد أم يقل كلما كبرت العينة)؟
- لو وجدت أن تباين مقدر غير متحيز أكبر من حد كرامير – راو، ماذا يمكنك أن تفعل؟

ثالثاً: تطبيق

الموقف 1:

تريد تقدير متوسط طول طلاب الجامعة. تستخدم متوسط العينة.

- كيف يمكنك التحقق مما إذا كان هذا المقدر "فعالاً" أم لا؟ صف الخطوات بالكلمات.

الموقف 2:

لديك توزيع برنولي (نسبة نجاح). تريد معرفة أقل تباين يمكن أن يصل إليه أي مقدر غير متحيز لنسبة النجاح.

- كيف تستخدم معلومة فيشر لحساب هذا الحد؟ اذكر الخطوات العامة.

الموقف 3:

- باحث وجد أن تباين مقدره غير المتحيز يساوي 0.25، بينما حد كرامير – راو لنفس المعلمة هو 0.10. -
- ماذا يستنتج الباحث؟ وماذا ينبغي أن يفعل؟

رابعاً: تقييم ذاتي سريع (اختر العبارة الصحيحة)

معلومة فيشر تقيس:

أ) مقدار التحيز في المقدر

(ب) كمية المعلومات التي تحملها العينة عن المعلمة

(ج) حجم العينة المناسب

(د) تباين المجتمع

متباينة كرامير – راو تعطي:

قيمة المعلمة الحقيقية

(ب) أقل تباين ممكن لأي مقدر (سواء متحيز أو غير متحيز)

(ج) أقل تباين ممكن للمقدرات غير المتحيزة

(د) أكبر تباين ممكن للمقدرات

المقدر "الفعال" هو مقدر:

(أ) متحيز وتباينه صغير

(ب) غير متحيز ويصل تباينه إلى حد كرامير – راو

(ج) غير متحيز وتباينه أكبر من حد كرامير – راو

(د) له تباين يساوي صفراً

شرط "مجال الدالة لا يعتمد على المعلمة" ينطبق على توزيع:

(أ) برنولي

(ب) بواسون

(ج) طبيعي

(د) منتظم $U(0, \theta)$

إذا زاد حجم العينة n ، فإن حد كرامير – راو:

(أ) يزداد

(ب) يقل

(ج) يبقى ثابتاً

(د) يتضاعف

المقدر الذي يحقق حد كرامير – راو يُعتبر:

(أ) أفضل مقدر متحيز

(ب) أفضل مقدر غير متحيز (أقل تباين)

(ج) أقل المقدرات كفاءة

(د) مقدر غير متنسق

في توزيع طبيعي بمعلومية التباين، مقدر المتوسط (متوسط العينة) هو مقدر:

أ) فعال

ب) غير فعال

ج) متحيز

د) غير متنسق

إذا كان تباين مقدر غير متحيز يساوي ضعف حد كرامير – راو، فإن كفاءته النسبية هي:

أ) 50%

ب) 100%

ج) 200%

د) لا يمكن حسابها

الفصل الثاني نظرية التقدير

الدرس السابع: التقدير المجالي (مجال الثقة للمتوسط في المجتمع، ومجال الثقة للنسبة في المجتمع)

المقدمة التعريفية:

سُيَمِّز بين التقدير النقطي والتقدير المجالي من حيث المفهوم والفائدة، سُنَحَد الصيغ المناسبة لبناء مجال ثقة لمتوسط المجتمع في ثلاث حالات: عندما يكون التباين معلوماً (بغض النظر عن حجم العينة)، وعندما يكون التباين مجهولاً مع عينة كبيرة ($n > 30$)، وعندما يكون التباين مجهولاً مع عينة صغيرة ($n < 30$) باستخدام توزيع ستودنت. سُنُطَبَق خطوات بناء مجال ثقة لنسبة المجتمع باستخدام التقريب الطبيعي (طريقة والد) مع التحقق من شروط صحة التقريب ($np \geq 5$) سُنُحَل أمثلة واقعية من استطلاعات الرأي، مراقبة الجودة، والدراسات الطبية لاستخراج مجالات الثقة وتفسيرها.

الهدف العام للدرس

أن يُثَقِّن الطالب بنهاية هذا الدرس بناء وتفسير مجال ثقة لمتوسط المجتمع (في حالات التباين المعلوم والمجهول) ولمعلمة النسبة في المجتمع، وذلك من خلال التمييز بين الحالات المختلفة، وحساب المجالات باستخدام الصيغ المناسبة (التوزيع الطبيعي أو توزيع ستودنت)، والتحقق من شروط الاستخدام، وتفسير النتائج في سياقات عملية.

الأهداف التفصيلية للدرس

- ❖ يُعَرِّف الطالب كلاً من: التقدير النقطي، التقدير المجالي (مجال الثقة)، مستوى الثقة.
- ❖ يُحَدِّد الطالب الصيغة المناسبة لبناء مجال ثقة لمتوسط المجتمع وفقاً لثلاث حالات: تباين المجتمع معلوم، تباين المجتمع مجهول $n > 30$ ، تباين المجتمع مجهول $n < 30$.
- ❖ يُطَبِّق الطالب خطوات بناء مجال ثقة للنسبة في المجتمع باستخدام طريقة والد (التقريب الطبيعي).
- ❖ يُحَلِّل الطالب جداول بيانات عينات (مثل أعمار التلاميذ، أوزان الأطفال، أعمار الشاشات، نسب الإصابة بالأمراض) لاستخراج مجالات ثقة بنسب 95% أو 90%، ويُفسر النتائج بلغة عملية.
- ❖ يُقَارِن الطالب بين اتساع مجال الثقة الناتج عن حالات مختلفة (حجم عينة صغير مقابل كبير، مستوى ثقة مرتفع مقابل منخفض، تباين معلوم مقابل مجهول).

مواضيع الدرس

- ❖ التقدير النقطي (Point Estimation)
- ❖ التقدير بمجال الثقة للمتوسط (Confidence Interval for Mean)
- ❖ التقدير بمجال الثقة للنسبة (Confidence Interval for Proportion)
- ❖ أهمية المتباينة وقيودها.

الوقت المتوقع: 4.5 ساعات بين 3 ساعات محاضرات و 1 ساعة و 30 دقيقة أعمال موجهة

تمهيد:

بعد أن تعرفنا في الدروس السابقة على التقدير النقطي الذي يقدم قيمة واحدة كتقدير لمعلمة المجتمع، أدركنا أن هذا التقدير وحده لا يعطينا فكرة عن مدى دقته أو قربته من القيمة الحقيقية. لذا ننتقل في هذا الدرس إلى التقدير المجالي (Interval Estimation) أو ما يُعرف بـ مجال الثقة (Confidence Interval)، وهو أسلوب إحصائي يقدم فترة (مجالاً) من القيم، مع درجة معينة من الثقة بأن المعلمة الحقيقية للمجتمع تقع ضمن هذه الفترة.

1- التقدير النقطي:

هو التقدير النقطي هو قيمة عددية واحدة لاحصاء من العينة الذي يمكننا من خلاله تقدير قيمة معالم المجتمع. إذن التقدير النقطي هو تقدير معالم المجتمع أو يصف توزيع المجتمع على ضوء معلومات حول المعاينة بنقطة واحدة.

فالمتوسط الحسابي للعينة $\bar{X}$ هو تقدير غير متحيز للمعلمة μ ، ومن خلال الانحراف المعياري للعينة σ يمكن تقدير الانحراف المعياري للمجتمع كذلك (يداوي محمد، دوة محمد، 2016، صفحة 75).

على سبيل المثال، فإنه إذا تم جمع بيانات من عينة من 500 شخص في المدينة ووجد أن المتوسط الحسابي لدخلهم السنوي هو 30,000 دولار، فيمكن استخدام هذا التقدير النقطي للتوصل إلى التقدير للمتوسط العام للدخل السنوي في المدينة بأنه 30,000 دولار.

الانحراف المعياري للعينة يعتبر مقدر متحيز أو يجب التقليل من قيمة التحيز من خلال إجراء تصحيح بسيط وذلك بإدخال معامل التصحيح (زهراري عفاف، 2020، صفحة 2).

إذا كان μ هو المتوسط الحسابي للمجتمع

إذا كان $\bar{X}$ هو المتوسط الحسابي لعينة بحجم n فإن:

$\bar{X}$ هو مقدر نقطي صحيح مطلق لـ μ ونرمز للمقدر بـ $\hat{\mu}$

حيث:

$$\hat{\mu} = \bar{X} = \frac{\sum X_i}{n}$$

مثال تطبيقي 1:

لدينا عينة من 10 تلاميذ من مدرسة عبد الحميد بن باديس، معدلاتهم مبينة في الجدول:

$(X_i - \bar{X})^2$	$(X_i - \bar{X})$	X_i
5.29	2.3	18
2.89	-1.7	14
0.49	-0.7	15
1.69	1.3	17
0.49	-0.7	15
0.09	0.3	16
0.09	0.3	16
2.89	-1.7	14
0.49	-0.7	15
1.69	1.3	17
16.1	المجموع	157

المطلوب: قدر نقطيا معالم المجتمع (معدلات طلبة مدرسة عبد الحميد بن باديس).

الحل:

إذا كان μ هو المتوسط الحسابي للمجتمع

إذا كان $\bar{x}$ هو المتوسط الحسابي لعينة بحجم n فإن:

$\bar{X}$ هو مقدر نقطي صحيح مطلق لـ μ ونرمز للمقدر بـ $\hat{\mu}$

حيث:

$$\hat{\mu} = \bar{X} = \frac{\sum X_i}{n}$$

$$= \frac{157}{10} = 15.7$$

إذا كان σ^2 هو التباين للمجتمع

$\hat{S}^2$ هو مقدر نقطي صحيح مطلق لـ σ^2

$$\hat{S}^2 = \frac{\sum (X_i - \bar{X})^2}{n - 1}$$

$$= \frac{16.10}{10 - 1} = 1.78$$

إذا كان σ هو الانحراف المعياري للمجتمع

$\hat{S}$ هو مقدر نقطي صحيح مطلق لـ σ

$$\hat{S} = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n - 1}} = 1.33$$

يجب ملاحظة أن المتوسط الحسابي قد يكون تقديراً نقطياً للقيمة الحقيقية.

مثال تطبيقي 2:

ليكن X متغير عشوائي يخضع للتوزيع الطبيعي وسطه الحسابي μ وتباينه σ^2

في عينة عشوائية حجمها $n=56$ تبين لنا أن:

$$\sum X_i = 644$$

$$\sum (X_i - \bar{X})^2 = 198$$

- قدر نقطياً معالم المجتمع.

- الحل:

إذا كان μ هو المتوسط الحسابي للمجتمع

إذا كان $\bar{X}$ هو المتوسط الحسابي لعينة بحجم n فإن:

$\bar{X}$ هو مقدر نقطي صحيح مطلق لـ μ ونرمز للمقدر بـ $\hat{\mu}$

حيث:

$$\hat{\mu} = \bar{X} = \frac{\sum X_i}{n}$$

$$= \frac{644}{56} = 11.5$$

إذا كان σ^2 هو التباين للمجتمع

$\hat{\sigma}^2$ هو مقدر نقطي صحيح مطلق لـ σ^2

$$\hat{\sigma}^2 = \frac{\sum (X_i - \bar{X})^2}{n - 1}$$

$$= \frac{198}{56 - 1} = 3.6$$

إذا كان σ هو الانحراف المعياري للمجتمع

$\hat{\sigma}$ هو مقدر نقطي صحيح مطلق لـ σ

$$\hat{\sigma} = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n - 1}} = 1.89$$

مثال تطبيقي 3:

نسبة الطلبة الراسبون في كلية الاقتصاد هي P ، وفي الفوج الرابع من الكلية وجدنا أن هناك 6 طلبة راسبون من 42 طالب.

-المطلوب: أوجد مقدرًا نقطيًا للنسبة P .

الحل:

إذا كان P هي نسبة الراسبين في الكلية (للمجتمع)

إذا كان $\bar{P}$ هي نسبة الراسبين لعينة بحجم n فإن:

$\bar{P}$ هو مقدر نقطي صحيح مطلق ل P ونرمز للمقدر بـ $\hat{P}$

حيث: $\bar{n} = 9$

و $n = 42$

$$\hat{P} = \bar{P} = \frac{\bar{n}}{n} =$$

$$\frac{9}{45} = 0.2 = 2\%$$

2- التقدير بمجال ثقة للمتوسط :

أ/ متوسط متغير عشوائي طبيعي تباينه معلوم:

ليكن X متغير عشوائي طبيعي متوسطه μ مجهول، و تباينه σ^2 معلوم، فإن مجال الثقة للمعلمة μ كما يلي
(Sharma, 2010, p. 309):

$$p\left(\bar{x} - \frac{\sigma}{\sqrt{n}}Z_{1-\frac{\alpha}{2}} \leq \mu \leq \bar{x} + \frac{\sigma}{\sqrt{n}}Z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

$(1 - \alpha)$ هو مقدار الثقة.

ب/ متوسط متغير عشوائي طبيعي تباينه غير معلوم (مجهول):

ليكن X متغير عشوائي متوسطه μ مجهول، و تباينه σ^2 مجهول، فإن مجال الثقة للمعلمة μ كما يلي (يداوي محمد، دوة محمد، 2016، الصفحات 77-88):

الحالة الأولى: $n > 30$

$$p\left(\bar{x} - \frac{\hat{S}}{\sqrt{n}}Z_{1-\frac{\alpha}{2}} \leq \mu \leq \bar{x} + \frac{\hat{S}}{\sqrt{n}}Z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

الحالة الثانية: $n < 30$

$$p\left(\bar{x} - \frac{\hat{S}}{\sqrt{n}} t_{\frac{\alpha}{2}} \leq \mu \leq \bar{x} + \frac{\hat{S}}{\sqrt{n}} t_{\frac{\alpha}{2}}\right) = 1 - \alpha$$

التقدير بمجال ثقة (أو فترة ثقة)

الحالة 1: تباين المجتمع معلوم $n > 30$ أو $n < 30$

$$p\left(\bar{x} - \frac{\sigma}{\sqrt{n}} Z_{1-\frac{\alpha}{2}} \leq \mu \leq \bar{x} + \frac{\sigma}{\sqrt{n}} Z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

الحالة 2: تباين المجتمع مجهول $n > 30$

$$p\left(\bar{x} - \frac{\hat{S}}{\sqrt{n}} Z_{1-\frac{\alpha}{2}} \leq \mu \leq \bar{x} + \frac{\hat{S}}{\sqrt{n}} Z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

الحالة 3: تباين المجتمع مجهول $n < 30$

$$p\left(\bar{x} - \frac{\hat{S}}{\sqrt{n}} t_{1-\frac{\alpha}{2}} \leq \mu \leq \bar{x} + \frac{\hat{S}}{\sqrt{n}} t_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

درجة الحرية: $t_{1-\frac{\alpha}{2}}^{v=n-1}$

مثال تطبيقي 1:

1/ لنفرض أنه أخذت عينة مكونة من 400 طفل من المدارس الصينية، من بين 100000 طفل متمدرس، وجد أن متوسط أوزانهم هو 27 كغ، وكان الانحراف المعياري للمجتمع 4 كغ.

- قم بتقدير متوسط أوزان التلاميذ في المجتمع المدرسي الصيني.

2/ من مصنع LG سحبت عينة عشوائية مكونة من 9 شاشات بمتوسط عمر 30000 ساعة وانحراف معياري قدره: 420 ساعة.

إذا علمت أن طول عمر الشاشات في المصنع يتبع التوزيع الطبيعي، قدر بمجال ثقة متوسط طول عمر الشاشات في المجتمع عند فترة ثقة 95%.

الحل:

1/ بتقدير متوسط أوزان التلاميذ في المجتمع المدرسي الصيني

$$\begin{cases} \sigma = 4 \text{ الانحراف المعياري للمجتمع معلوم} \\ n = 400 > 30 \end{cases}$$

فإن توزيع المعاينة يتبع التوزيع الطبيعي:

$$p\left(\bar{x} - \frac{\sigma}{\sqrt{n}} Z_{1-\frac{\alpha}{2}} \leq \mu \leq \bar{x} + \frac{\sigma}{\sqrt{n}} Z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$\alpha = 0.1$$

$$Z_{1-\frac{\alpha}{2}} = Z_{1-\frac{0.1}{2}} = Z_{0.9500} = 1.65$$

$$p\left(27 - \frac{4}{\sqrt{400}} 1.65 \leq \mu \leq 27 + \frac{4}{\sqrt{400}} 1.65\right) = 95\%$$

$$p(26.67 \leq \mu \leq 27.33) = 95\%$$

2/ بتقدير متوسط عمر الشاشات في المجتمع (المصنع)

$$\begin{cases} \sigma = ? \text{ الانحراف المعياري مجهول للمجتمع} \\ n = 9 < 30 \end{cases}$$

فإن توزيع المعاينة يتبع توزيع ستودنت:

$$p\left(\bar{x} - \frac{\hat{S}}{\sqrt{n}} t_{1-\frac{\alpha}{2}} \leq \mu \leq \bar{x} + \frac{\hat{S}}{\sqrt{n}} t_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

أولا حساب مقدرة الانحراف للمجتمع: $\hat{S}$

$$\hat{S} = \sqrt{\hat{S}^2}$$

$$\text{و } \hat{S}^2 = \frac{\sum (X_i - \bar{X})^2}{n - 1}$$

$$\sigma_{\bar{x}}^2 = \frac{\sum (X_i - \bar{X})^2}{n} \rightarrow \sum (X_i - \bar{X})^2$$

$$= n \cdot \sigma_{\bar{x}}^2$$

$$= 9 \times 420^2 = 1587600$$

$$\boxed{\sum (X_i - \bar{X})^2 = 1587600}$$

$$\hat{S}^2 = \frac{\sum (X_i - \bar{X})^2}{n - 1}$$

$$= \frac{1587600}{8} = 198450$$

$$\hat{S} = \sqrt{\hat{S}^2}$$

$$= \sqrt{198450} = 445.47$$

ثانيا حساب $T_{1-\frac{\alpha}{2}}$

$t_{0.975}^{v=8}$ تساوي 2.306

$$p\left(\bar{x} - \frac{\hat{S}}{\sqrt{n}} t_{1-\frac{\alpha}{2}} \leq \mu \leq \bar{x} + \frac{\hat{S}}{\sqrt{n}} t_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$p(29657.58 \leq \mu \leq 30342.42) = 95\%$$

2- التقدير بمجال ثقة للنسبة :

بعد أن تعرفنا في الدروس السابقة على التقدير النقطي (Point Estimation) ، الذي يقدم قيمة وحيدة كتقدير لمعلمة المجتمع (مثل تقدير نسبة النجاح p بقيمة واحدة من العينة)، ندرك أن هذا التقدير وحده لا يعطينا فكرة عن مدى دقته أو قربيه من القيمة الحقيقية في المجتمع. فلو قدرنا أن نسبة المؤيدين لقضية ما هي 60% من عينة حجمها 100 ، فهل هذه النسبة قريبة من النسبة الحقيقية؟ وهل يمكن الوثوق بها؟

هنا تأتي أهمية التقدير بفواصل الثقة (Interval Estimation) أو ما يُعرف بـ مجال الثقة (Confidence Interval). فبدلاً من تقديم قيمة واحدة، نقدم فترة (مجالاً) (من القيم، ونكون على درجة معينة من الثقة (مثل 95% أو 99%) بأن المعلمة الحقيقية للمجتمع تقع ضمن هذه الفترة.

عندما تكون المعلمة محل الاهتمام هي نسبة النجاح في المجتمع (p) مثل نسبة المصوتين لمرشح، نسبة المنتجات المعيبة، نسبة المرضى المتعافين...، فإننا نستخدم بيانات العينة لحساب نسبة النجاح في العينة ($\bar{p}$)، ثم نبني حولها مجالاً يعتمد على: (Sharma, 2010, pp. 314-315)

– حجم العينة (n)

– مستوى الثقة المطلوب -مثلاً 95%

التوزيع الاحتمالي المناسب (عادة التوزيع الطبيعي بموجب نظرية النهاية المركزية، بشرط توفر شروط $np \geq 5$)

يمثل هذا المجال نطاقاً من القيم المحتملة للنسبة الحقيقية، ويمكننا القول مثلاً: "بتقنة 95%"، فإن نسبة المؤيدين في المجتمع تتراوح بين 52% و 68%". هذا النوع من التقدير أكثر إفادة في اتخاذ القرارات الإحصائية، لأنه يوضح هامش الخطأ المحتمل ودرجة عدم اليقين المرتبطة بالتقدير.

في هذا الجزء من الدرس، سنتعلم كيفية بناء مجال ثقة للنسبة، والصيغ الرياضية المستخدمة، والشروط الواجب توافرها، وسنطبق ذلك على أمثلة واقعية من استطلاعات الرأي ومراقبة الجودة والدراسات الطبية.

عندما نريد تقدير نسبة النجاح في المجتمع p مثل نسبة المصوتين، نسبة المعيب، نسبة الشفاء، نستخدم نسبة النجاح في العينة ($\bar{p}$) ونبني فاصل الثقة كما يلي:

فاصل الثقة التقريبي (باستخدام التوزيع الطبيعي – طريقة والد (Wald))

إذا توفرت شروط التقريب الطبيعي ($np \geq 5$) فإن فاصل الثقة التقريبي لمستوى ثقة $1-\alpha$ يُعطى بالصيغة: (Prem & Christopher, 2010, p. 374):

$$p \left(\bar{p} - (\sigma_{\bar{p}} Z_{1-\frac{\alpha}{2}}) \leq p \leq \bar{p} + (\sigma_{\bar{p}} Z_{1-\frac{\alpha}{2}}) \right) = 1 - \alpha$$

صيغة ويلسون – (Wilson Score Interval) للعينات الصغيرة أو الحالات الحدودية

هذه الصيغة تعطي نتائج أفضل عندما تكون n صغيرة أو $\bar{p}$ قريب من 0 أو 1 (Sharma, 2010, pp. 314-315):

$$\frac{\bar{p} + \frac{z^2}{2n} \mp Z \sqrt{\frac{\bar{p}(1-\bar{p})}{n} + \frac{z^2}{4n^2}}}{1 + \frac{z^2}{n}}$$

مثال تطبيقي 2:

في مدينة صناعية معينة، وفي عينة من 120 شخص تبين أن عدد المصابين بالربو هو 55.

- شكل مجال ثقة لنسبة الإصابة بالربو في المدينة الصناعية عند مستوى ثقة 95%

إذا كان P هي نسبة المصابين في المدينة (للمجتمع)

إذا كان $\bar{X}$ هي نسبة المصابين في العينة بحجم $n=120$ فإن:

$$p \left(\bar{p} - (\sigma_{\bar{p}} Z_{1-\frac{\alpha}{2}}) \leq p \leq \bar{p} + (\sigma_{\bar{p}} Z_{1-\frac{\alpha}{2}}) \right) = 1 - \alpha$$

حيث: $\bar{n} = 55$ و $n = 120$

$$\bar{P} = \frac{\bar{n}}{n} = \frac{55}{120} = 0.45 = 45\%$$

$$\bar{q} = 1 - \bar{P} = 0.55$$

$$p \left(\bar{p} - (\sigma_{\bar{p}} Z_{1-\frac{\alpha}{2}}) \leq p \leq \bar{p} + (\sigma_{\bar{p}} Z_{1-\frac{\alpha}{2}}) \right) = 1 - \alpha$$

$$\sigma_{\bar{p}} = \sqrt{\frac{\bar{p}\bar{q}}{n}} = 0.044$$

$$1 - \alpha = 0.95$$

$$\rightarrow \alpha = 0.05$$

$$Z_{1-\frac{\alpha}{2}} = Z_{1-\frac{0.05}{2}}$$

$$= Z_{0.995} = 1.96$$

بالتعويض نجد:

$$p\left(\bar{p} - (\sigma_{\bar{p}} Z_{1-\frac{\alpha}{2}}) \leq p \leq \bar{p} + (\sigma_{\bar{p}} Z_{1-\frac{\alpha}{2}})\right) = 1 - \alpha$$

$$p(0.45 - (0.044 \times 1.96) \leq p \leq 0.45 + (0.044 \times 1.96)) = 95\%$$

$$p(0.3637 \leq p \leq 0.5362) = 95\%$$

وهو مجال الثقة المطلوب.

التقييم الذاتي للدرس السابع: التقدير المجالي (مجال الثقة للمتوسط والنسبة)

أولاً: أسئلة للفهم (أجب بكلمة أو جملة قصيرة)

- ما الفرق بين التقدير النقطي والتقدير المجالي؟
- ماذا يعني "مستوى ثقة 95%" في مجال الثقة؟
- متى نستخدم توزيع "ستيوذنت (t)" بدلاً من التوزيع الطبيعي (Z) عند تقدير متوسط المجتمع؟
- اذكر شرطاً واحداً لاستخدام التقريب الطبيعي عند بناء مجال ثقة لنسبة المجتمع.
- ماذا يحدث لعرض مجال الثقة (اتساعه) إذا زاد حجم العينة مع بقاء مستوى الثقة نفسه؟

ثانياً: أسئلة تحليلية

- لماذا يعتبر التقدير المجالي أكثر فائدة من التقدير النقطي في اتخاذ القرارات الإحصائية؟
- لو كان لديك عينة صغيرة (عددها أقل من 30) وتريد تقدير متوسط مجتمع طبيعي ولكنك لا تعرف تباين المجتمع، أي توزيع تستخدم؟ ولماذا؟
- كيف يتغير مجال الثقة للنسبة إذا رفعت مستوى الثقة من 90% إلى 99%؟ (هل يتسع أم يضيق؟)
- في أي حالة يكون مجال الثقة للمتوسط أضيق: عندما يكون التباين معلوماً أم عندما يكون مجهولاً (مع نفس حجم العينة ومستوى الثقة)؟ ولماذا؟
- لماذا نستخدم القيمة الحرجة من توزيع t بدلاً من Z عندما يكون حجم العينة أقل من 30 والتباين مجهولاً؟

ثالثاً: تطبيق

التمرين 1:

يريد باحث تقدير متوسط الإنفاق الشهري للأسر في مدينة. لديه عينة عشوائية من 400 أسرة، ويعرف الانحراف المعياري للإنفاق في المجتمع من دراسات سابقة. صف الخطوات التي سيتبعها لبناء مجال ثقة 95% للمتوسط.

التمرين 2:

في مصنع للأجهزة الكهربائية، سحبت عينة عشوائية من 12 جهازاً فقط (لأن الإنتاج محدود). لا يعرف الباحث تباين عمر الأجهزة في المصنع، لكنه يعرف أن الأعمار تتبع التوزيع الطبيعي. ماذا يجب أن يستخدم بدلاً من Z؟ وما الذي يحتاج إليه من جدول إحصائي؟

التمرين 3:

في استطلاع رأي على 200 شخص، وجد أن 120 يؤيدون قراراً معيناً. هل تنطبق شروط استخدام التقريب الطبيعي لبناء مجال ثقة للنسبة؟ تحقق من الشرطين المطلوبين (عدد المؤيدين وعدد المعارضين كافيان) دون حساب.

رابعاً: تقييم ذاتي سريع (اختر العبارة الصحيحة)

التقدير النقطي لمتوسط المجتمع هو:

(أ) قيمة واحدة تحسب من العينة

(ب) فترة من القيم

(ج) مستوى ثقة

(د) خطأ معياري

مستوى الثقة 95% يعني:

(أ) 95% من بيانات العينة تقع في المجال

(ب) 95% من مجالات الثقة المبنية من عينات متكررة تحتوي على المعلمة الحقيقية

(ج) احتمال أن يكون التقدير النقطي صحيحاً هو 95%

(د) خطأ المعاينة هو 5%

عند تقدير متوسط مجتمع والتباين مجهول، وحجم العينة 25، نستخدم:

(أ) التوزيع الطبيعي (Z)

(ب) توزيع مربع كاي

(ج) توزيع ستودنت (t) بدرجة حرية 24

(د) توزيع F

إذا زاد حجم العينة مع بقاء مستوى الثقة ثابتاً، فإن عرض مجال الثقة:

(أ) يزداد

(ب) يقل

(ج) لا يتغير

(د) يتضاعف

شروط استخدام التقريب الطبيعي لمجال ثقة النسبة هو:

(أ) حجم العينة أكبر من 30 فقط

(ب) عدد النجاحات وعدد الإخفاقات في العينة كلاهما أكبر من أو يساوي 5

(ج) المجتمع طبيعي التوزيع

(د) مستوى الثقة 95%

القيمة الحرجة Z لمستوى ثقة 90% تساوي تقريباً:

أ) 1.65

ب) 1.96

ج) 2.58

د) 2.33

أي مما يلي يؤثر على اتساع مجال الثقة للمتوسط؟

أ) حجم العينة

ب) التباين (أو الانحراف المعياري)

ج) مستوى الثقة

د) كل ما سبق

مجال الثقة الذي يحتوي على المعلمة الحقيقية باحتمال كبير يُفضل أن يكون:

أ) واسعاً جداً

ب) ضيقاً قدر الإمكان (بدقة مقبولة)

ج) قيمته صفر

د) معتمداً على التقدير النقطي فقط

الفصل الثاني نظرية التقدير

الدرس الثامن: التقدير المجالي للتباين، ونسبة تباين مجتمع طبيعي

المقدمة التعريفية:

بعد أن تعلمنا في الدروس السابقة كيفية بناء مجالات الثقة للمتوسط والنسبة في مجتمع واحد، ننتقل اليوم إلى موضوع جديد ومهم: تقدير تباين المجتمع، ونسبة تباين مجتمعين طبيعيين، ففي كثير من التطبيقات العملية، لا نهتم فقط بمتوسط المجتمع، بل بمدى تشتت القيم حول المتوسط (أي التباين)، على سبيل المثال: نريد تقدير تباين أعمار بطاريات السيارات، أو مقارنة مدى تجانس منتجين من مصنعين مختلفين.

في هذا الدرس، سنعرّف مفهوم مجال الثقة لتباين مجتمع طبيعي باستخدام توزيع مربع كاي- (Chi-Square) سنحدد خطوات بناء مجال الثقة للتباين وللانحراف المعياري، سنطبق طريقة بناء مجال الثقة لنسبة تباين مجتمعين طبيعيين مستقلين باستخدام توزيع F ونحل العلاقة بين درجة الحرية ومستوى الثقة وعرض المجال.

الهدف العام للدرس

أن يتقن الطالب بنهاية هذا الدرس بناء وتفسير مجال ثقة لتباين مجتمع طبيعي واحد، ولمجال ثقة لنسبة تباين مجتمعين طبيعيين مستقلين، وذلك باستخدام توزيعي مربع كاي و F ، مع التحقق من شروط التطبيق (طبيعية المجتمعات، استقلال العينات)، وحساب المجالات من بيانات معطاة، وتفسير النتائج في سياقات عملية.

الأهداف التفصيلية للدرس

- ❖ يُعرّف الطالب المصطلحات: مجال ثقة للتباين، مجال ثقة لنسبة التباينين.
- ❖ يُطبق الطالب خطوات حساب مجال ثقة لانحراف معياري (بأخذ الجذر التربيعي للحدود).
- ❖ يحلل الطالب بيانات عينتين مستقلتين من مجتمعين طبيعيين لبناء مجال ثقة لنسبة تباينيهما باستخدام توزيع F .
- ❖ يُقارن الطالب بين مجالات الثقة للتباين في حالات: عينة صغيرة vs كبيرة، ومستويات ثقة مختلفة، ويقيم مدى دقة تقدير التباين مقارنة بتقدير المتوسط.

مواضيع الدرس

- ❖ مجال الثقة لتباين مجتمع طبيعي واحد؛
- ❖ مجال الثقة لنسبة تباين مجتمعين طبيعيين مستقلين؛
- ❖ اشتقاق مجال الثقة لنسبة التباينين.

الوقت المتوقع: 4.5 ساعات بين 3 ساعات محاضرات و 1 سا و 30 حصة أعمال موجهة

تمهيد:**1- مجال الثقة لتباين مجتمع طبيعي واحد**

نرغب في تقدير التباين (σ^2)، (أو الانحراف المعياري σ) بمجال ثقة $(1-\alpha)$ ، من تعريف توزيع مربع كاي (Sharma, 2010, p. 397):

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$$

$$P\left(\chi_{1-\frac{\alpha}{2}, n-1}^2 \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi_{\frac{\alpha}{2}, n-1}^2\right) = 1 - \alpha$$

حيث $\chi_{1-\frac{\alpha}{2}, n-1}^2$ هي القيمة التي يترك توزيع مربع كاي على يمينها مساحة $\frac{\alpha}{2}$ (أي أقل من توزيع كاي تربيع)، بينما $\chi_{\frac{\alpha}{2}, n-1}^2$ هي القيمة التي يترك على يمينها مساحة $\frac{\alpha}{2}$ (1) ويطلق عليها أحياناً "القيمة الحرجة العليا".

بإعادة ترتيب المتباينة نحصل على مجال الثقة لـ σ^2 :

$$\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}, n-1}^2} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}, n-1}^2}$$

بالنسبة للانحراف المعياري σ ، نأخذ الجذر التربيعي للحدود:

$$\sqrt{\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}, n-1}^2}} \leq \sigma \leq \sqrt{\frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}, n-1}^2}}$$

مثال تطبيقي:

المعطيات: عينة عشوائية من 10 قطع منتجة من آلة، تباين العينة $S^2 = 15.2$ افترض أن الإنتاج يتبع توزيعاً طبيعياً، أوجد مجال ثقة 95% لتباين المجتمع σ^2 .

الحل:

$$n=10$$

$$V=n-1=9$$

$$1-\alpha=0.95$$

$$\alpha=0.05$$

$$\alpha/2=0.025$$

من جدول توزيع مربع كاي:

القيمة الحرجة العليا:

$$\chi_{\frac{\alpha}{2}, n-1}^2 = \chi_{0.025, 9}^2 = 19.023$$

القيمة الحرجة السفلى:

$$\chi_{\frac{\alpha}{2}-1, n-1}^2 = \chi_{0.975, 9}^2 = 2.700$$

نطبق المعادلة:

$$\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}, n-1}^2} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}, n-1}^2}$$

$$\frac{(10-1)15.2}{19.023} \leq \sigma^2 \leq \frac{(10-1)2}{2.700}$$

$$\frac{(10-1)15.2}{19.023} \leq \sigma^2 \leq \frac{(10-1)2}{2.700}$$

$$\frac{136.8}{19.023} \leq \sigma^2 \leq \frac{136.8}{2.700}$$

$$50.667 \leq \sigma^2 \leq 7.190$$

تقدير الانحراف المعياري:

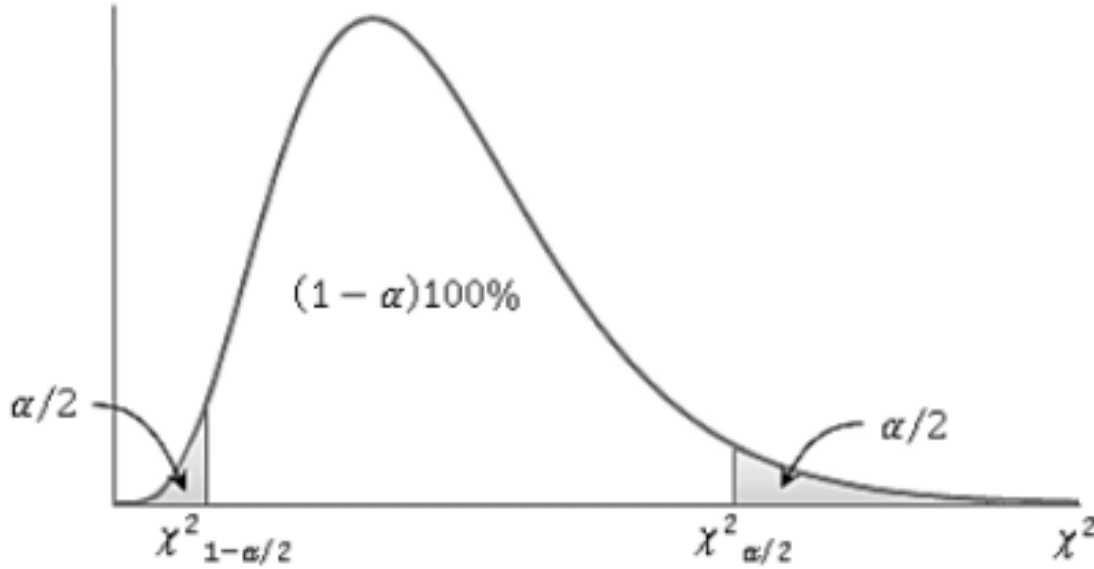
$$\sqrt{\frac{(n-1)S^2}{\chi_{\frac{\alpha}{2}, n-1}^2}} \leq \sigma \leq \sqrt{\frac{(n-1)S^2}{\chi_{1-\frac{\alpha}{2}, n-1}^2}}$$

$$\sqrt{50.667} \leq \sigma \leq \sqrt{7.190}$$

$$2.68 \leq \sigma \leq 7.12$$

وهو تقدير الانحراف المعياري للمجتمع بمجال ثقة 95%.

الشكل رقم 2: التمثيل البياني لتوزيع مربع كاي.



المصدر: (Sharma, 2010, p. 382).

2- مجال ثقة لنسبة تباين مجتمعين طبيعيين مستقلين

لنفرض أن لدينا عينتين مستقلتين من مجتمعين طبيعيين:

إذا كان (S_1^2) هو تباين عينه عشوائية حجمها n_1 مسحوبة من مجتمع طبيعي $N(\mu_1, \sigma_1)$ ، وكان (S_2^2) هو تباين عينه عشوائية أخرى حجمها n_2 ومسحوبة من مجتمع طبيعي آخر $N(\mu_2, \sigma_2)$ وكانت العينتان مستقلتان (Sharma, 2010, p. 367):

نعلم أن:

$$\frac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$$

حيث أن $F(n_1 - 1, n_2 - 1)$ تسمى بتوزيع F بدرجتي الحرية $n_1 - 1$ و $n_2 - 1$ و دالة الكثافة

الإحتمالية للمتغير y الذي يخضع لتوزيع F بدرجتي الحرية v_1, v_2 :

$$f(y) = \frac{y^{\frac{v_1}{2}-1}}{(v_1 y + v_2)^{\frac{v_1+v_2}{2}}} , \quad y > 0$$

حيث F هو توزيع فيشر - سنيديكور.

نريد بناء مجال ثقة لنسبة التباينين σ_1 و σ_2 نبدأ بالعلاقة:

$$P \left(F_{1-\frac{\alpha}{2}, v_1, v_2} \leq \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \leq F_{\frac{\alpha}{2}, v_1, v_2} \right) = 1 - \alpha$$

حيث أن $F(n_1 - 1, n_2 - 1)$ تسمى بتوزيع F بدرجتي الحرية $n_1 - 1$ و $n_2 - 1$ و دالة الكثافة

الاحتمالية للمتغير y الذي يخضع لتوزيع F بدرجتي الحرية v_1, v_2 :

نعكس المتباينة لتصبح بالنسبة لـ $\frac{\sigma_1^2}{\sigma_2^2}$:

$$P \left(\frac{S_1^2}{S_2^2} * \frac{1}{F_{\frac{\alpha}{2}, v_1, v_2}} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{S_1^2}{S_2^2} * F_{\frac{\alpha}{2}, v_1, v_2} \right) = 1 - \alpha$$

مثال تطبيقي:

المعطيات:

عينتان مستقلتان من إنتاج مصنعين للبطاريات:

$$S_1^2 = 22.5, n_1 = 12 \quad \text{المصنع A}$$

$$S_2^2 = 9.8, n_2 = 10 \quad \text{المصنع B}$$

بافتراض التوزيع الطبيعي، أوجد مجال ثقة 90% لنسبة التباين $\frac{\sigma_1^2}{\sigma_2^2}$

الحل:

$$v_1 = 11$$

$$v_2 = 9$$

$$\alpha = 0.90$$

$$1 - \alpha = 0.10$$

$$1 - \frac{\alpha}{2} = 0.05$$

من جدول توزيع F

القيمة الحرجة العليا لبسط 11 ومقام 9:

$$F_{\frac{\alpha}{2}, v_1, v_2} = F_{0.05, 11, 9} = 2.90$$

لعكس الدرجات:

$$F_{\frac{\alpha}{2}, v_1, v_2} = F_{0.05, 9, 11} = 2.59$$

نسبة التباين للعينة:

$$\frac{S_1^2}{S_2^2} = \frac{22.5}{9.8} = 2.2959$$

الطرف الأدنى:

$$\frac{S_1^2}{S_2^2} * \frac{1}{F_{\frac{\alpha}{2}, v_1, v_2}} = 0.7917$$

$$\frac{S_1^2}{S_2^2} * F_{\frac{\alpha}{2}, v_1, v_2} = 5.946$$

$$P\left(0.7917 \leq \frac{\sigma_1^2}{\sigma_2^2} \leq 5.946\right) = 0.10$$

تفسير: نثق بنسبة 90% أن تباين عمر بطاريات المصنع A يتراوح بين 0.79 و 5.95 ضعف تباين المصنع B. بما أن المجال يحتوي على 1، لا يمكننا رفض فرض تساوي التباينين عند مستوى دلالة 0.10.

التقييم الذاتي للدرس الثامن

أولاً: أسئلة للفهم

- ما اسم التوزيع الذي نستخدمه لبناء مجال ثقة لتباين مجتمع طبيعي واحد؟
- ماذا تمثل "درجات الحرية" في هذا السياق؟
- لماذا نستخدم توزيع F عند مقارنة تبايني مجتمعين طبيعيين؟
- هل مجال الثقة للتباين متماثل حول تقدير العينة؟
- إذا كان مجال الثقة لنسبة تباينين يحتوي على الرقم 1، ماذا يمكن أن نستنتج؟

ثانياً: أسئلة تحليلية

- كيف يتغير عرض مجال الثقة للتباين إذا زاد مستوى الثقة من 90% إلى 99% (مع بقاء حجم العينة ثابتاً)؟
- لماذا نستخدم القيم الحرجة من جدول مربع كاي بشكل معكوس (الكبير في المقام والصغير في الطرف الآخر)؟
- في أي حالة يكون مجال الثقة للانحراف المعياري مفيداً أكثر من مجال الثقة للتباين؟

ثالثاً: تمرين تطبيقي

- لديك عينة من 20 قياساً لدرجة حرارة فرن، وتريد تقدير تباين درجة الحرارة (مدى استقرار الفرن). تباين العينة المحسوب هو 4.5 درجة مربعة، قم باتباع الخطوات لبناء مجال ثقة 95% للتباين الحقيقي.
- إذا كانت القيم الحرجة لمربع كاي هي: $\chi^2_{0.025,9} = 19.02$ و $\chi^2_{0.975,9} = 2.70$ وحجم العينة $n=10$ و $S^2 = 20$ احسب الطرف الأعلى لمجال الثقة 95% للتباين.

رابعاً: اختيار الإجابة الصحيحة

لحساب مجال ثقة لتباين مجتمع طبيعي، نستخدم توزيع:

- A. Normal
- B. t-Student
- C. Chi-Square
- D. F

عند مقارنة تبايني مجتمعين طبيعيين مستقلين، إحصاء الاختبار يتبع توزيع:

- E. Normal

F. Chi-Square

G. F

H. Poisson

كلما زاد حجم العينة، فإن مجال الثقة للتباين:

أ) يصبح أوسع (عرض أكبر)

ب) يصبح أضيق

ج) لا يتغير عرضه

د) يتجه نحو الصفر

مجال ثقة لنسبة تباينين يحتوي على 1 يعني:

أ) التباين الأول يساوي الثاني قطعاً

ب) لا توجد فروق معنوية بين التباينين عند مستوى الثقة المستخدم

ج) التباين الأول أكبر من الثاني

د) العينات غير كافية

الفصل الثاني نظرية التقدير

الدرس التاسع: التقدير مجال الثقة للفرق بين متوسطي مجتمعين والفرق بين نسبتي

المقدمة التعريفية:

في هذا الدرس، سنُميّز بين حالات بناء فاصل الثقة للفرق بين متوسطين وفقاً لمعرفة تباين المجتمعين وحجم العينتين، سنطبق الصيغ المناسبة لكل حالة: عندما يكون التباينان معلومين، أو مجهولين مع عينات كبيرة، أو مجهولين مع عينات صغيرة (باستخدام توزيع t)، ونحلل بيانات عينتين مستقلتين لبناء فاصل ثقة للفرق بين نسبتي باستخدام التقريب الطبيعي. سنقارن بين اتساع الفترات الناتجة في الحالات المختلفة وعلاقتها بحجم العينة ومستوى الثقة. وأخيراً، نُقيّم مدى ملائمة كل طريقة وفقاً لمعطيات المسألة.

الهدف العام للدرس

أن يُتقن الطالب بنهاية هذا الدرس بناء وتفسير فاصل ثقة للفرق بين متوسطي مجتمعين (في حالات: تباينان معلومان، تباينان مجهولان مع عينات كبيرة، تباينان مجهولان مع عينات صغيرة) وللفرق بين نسبتي مجتمعين (للعينات الكبيرة)، وذلك باستخدام التوزيع الطبيعي أو توزيع t ، وبالتطبيق على أمثلة واقعية.

الأهداف التفصيلية

- ❖ يُعرّف الطالب الفرق بين متوسطي مجتمعين.
- ❖ يُطبق الطالب خطوات الصيغة المناسبة لبناء فاصل الثقة للفرق بين متوسطين.
- ❖ يُقارن الطالب بين قارن الطالب بين فترات الثقة الناتجة عن الحالات المختلفة.

مواضيع الدرس

- ❖ لتقدير مجال ثقة للفرق بين متوسطي مجتمعين؛
- ❖ التقدير بمجال ثقة للفرق بين نسبتي مجتمعين؛

الوقت المتوقع: 4.5 ساعات بين 3 ساعات محاضرات و 1 سا و 30د حصة أعمال موجهة

تمهيد:

في كثير من الدراسات الإحصائية، لا نكتفي بتقدير معلمة مجتمع واحد، بل نحتاج إلى مقارنة مجتمعين. هل متوسط إنتاجية العمال في مصنع (أ) يختلف عن المتوسط في مصنع (ب)؟ هل نسبة المصابين بمرض في مدينة (س) أقل منها في مدينة (ص)؟ للإجابة عن هذه الأسئلة، نستخدم التقدير بفواصل الثقة للفرق بين متوسطين أو الفرق بين نسبتيين.

1- التقدير بمجال ثقة للفرق بين متوسطي عيني مجتمعين:

إذا كان لدينا مجتمعين الأول له متوسط μ_1 وانحراف معياري σ_1 ، والثاني له متوسط μ_2 وانحراف معياري σ_2 ، فإن تقدير الفرق بين المتوسطين $(\mu_1 - \mu_2)$ هو الاحصاء $(\bar{x}_1 - \bar{x}_2)$ ، أي الفرق بين متوسطي عينتين من المجتمعين، حجم كل عينة n_1 و n_2 على التوالي (ثروت، 2023، الصفحات 250-254).¹ حيث نكون أمام الحالات التالية:

الحالة 1: تباين المجتمعين معومين σ_1^2, σ_2^2

$$p\left((\bar{x}_1 - \bar{x}_2) - \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} Z_{1-\frac{\alpha}{2}} \leq (\mu_1 - \mu_2) \leq (\bar{x}_1 - \bar{x}_2) + \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} Z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

أو:

$$p\left((\bar{x}_1 - \bar{x}_2) - \sigma_{\bar{x}_1 - \bar{x}_2} Z_{1-\frac{\alpha}{2}} \leq (\mu_1 - \mu_2) \leq (\bar{x}_1 - \bar{x}_2) + \sigma_{\bar{x}_1 - \bar{x}_2} Z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

الحالة 2: تباين المجتمعين مجهولين σ_1^2, σ_2^2 و $n_1, n_2 \geq 30$

$$p\left((\bar{x}_1 - \bar{x}_2) - \sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}} Z_{1-\frac{\alpha}{2}} \leq (\mu_1 - \mu_2) \leq (\bar{x}_1 - \bar{x}_2) + \sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}} Z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

أو:

¹ ثروت محمد عبد المنعم، مدخل حديث للإحصاء والاستدلال، الطبعة 3، دار كعيبان، الرياض، المملكة العربية السعودية، ص 250-254.

$$p \left((\bar{x}_1 - \bar{x}_2) - S^* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} Z_{1-\frac{\alpha}{2}} \leq (\mu_1 - \mu_2) \leq (\bar{x}_1 - \bar{x}_2) + S^* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} Z_{1-\frac{\alpha}{2}} \right) = 1 - \alpha$$

الحالة 3: تباين المجتمعين مجهولين σ_1^2, σ_2^2 و $n_1, n_2 < 30$

$$p \left((\bar{x}_1 - \bar{x}_2) - \sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}} t_{1-\frac{\alpha}{2}} \leq (\mu_1 - \mu_2) \leq (\bar{x}_1 - \bar{x}_2) + \sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}} t_{1-\frac{\alpha}{2}} \right) = 1 - \alpha$$

$$p \left((\bar{x}_1 - \bar{x}_2) - S^* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{1-\frac{\alpha}{2}} \leq (\mu_1 - \mu_2) \leq (\bar{x}_1 - \bar{x}_2) + S^* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{1-\frac{\alpha}{2}} \right) = 1 - \alpha$$

درجة الحرية: $\nu = n_1 + n_2 - 2$

S^{*2} مقدرة تباين الفرق بين متوسطي مجتمعين

$$S^{*2} = \frac{\sum (x_{i1} - \bar{x}_1)^2 + \sum (x_{i2} - \bar{x}_2)^2}{n_1 + n_2 - 2}$$

S^* مقدرة الانحراف المعياري للفرق بين متوسطي مجتمعين

$$S^* = \sqrt{S^{*2}}$$

مثال تطبيقي 1:

اخترنا عينة من الطالبات بحجم 75 طالبة فوجنا ان متوسط معدلاتهن 16 بانحراف معياري 1.25
واخترنا عينة اخرى بحجم 50 طالبا فتحصلنا على متوسط معدلاتهم 15.5 بانحراف معياري 0.75
المطلوب:

- أوجد تقديرا للفرق بين متوسطات مجتمع معدلات الطالبات ومجتمع معدلات الطلبة بمستوى ثقة 95%.

$$\left\{ \begin{array}{l} \text{تباين المجتمعين مجهولين} \\ n_1, n_2 > 30 \end{array} \right.$$

فإننا نستخدم مقدرة تباين الفرق بين متوسطي مجتمعين

$$\sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}}$$

بدلا من:

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

بما أن $n_1, n_2 > 30$ فان توزيع الفرق طبيعي: حيث:

$$p \left((\bar{x}_1 - \bar{x}_2) - \sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}} Z_{1-\frac{\alpha}{2}} \leq (\mu_1 - \mu_2) \leq (\bar{x}_1 - \bar{x}_2) + \sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}} Z_{1-\frac{\alpha}{2}} \right) = 1 - \alpha$$

حساب $Z_{1-\frac{\alpha}{2}}$:

لدينا

$$1 - \alpha = 0.95 \rightarrow \alpha = 0.05$$

ومنه:

$$Z_{1-\frac{\alpha}{2}} = Z_{1-\frac{0.05}{2}} = Z_{0.9750} = 1.96$$

حساب $\hat{S}_2^2$ و $\hat{S}_1^2$:

$$\hat{S}^2 = \frac{\sum (X_i - \bar{X})^2}{n - 1}$$

$$\sigma_{x1}^2 = \frac{\sum (X_{i1} - \bar{X}_1)^2}{n_1} \rightarrow \sum (X_{i1} - \bar{X}_1)^2$$

$$= n_1 \cdot \sigma_{x1}^2 = 75 \times 1.25^2 = 117.1875$$

$$\boxed{\sum (X_{i1} - \bar{X}_1)^2 = 117.1875}$$

$$\hat{S}_1^2 = \frac{\sum (X_{i1} - \bar{X}_1)^2}{n_1 - 1}$$

$$= \frac{117.1875}{74} = 1.5836$$

$$\sigma_{x2}^2 = \frac{\sum (X_{i2} - \bar{X}_2)^2}{n_2} \rightarrow \sum (X_{i2} - \bar{X}_2)^2$$

$$= n_2 \cdot \sigma_{x2}^2 = 50 \times 0.75^2 = 28.125$$

$$\boxed{\sum (X_{i2} - \bar{X}_2)^2 = 28.125}$$

$$\hat{S}_2^2 = \frac{\sum (X_{i2} - \bar{X}_2)^2}{n_2 - 1} = \frac{28.125}{49} = 0.5739$$

$$p \left((16 - 15.5) - \sqrt{\frac{1.5836}{75} + \frac{0.5739}{50}} 1.96 \leq (\mu_1 - \mu_2) \right.$$

$$\left. \leq (16 - 15.5) + \sqrt{\frac{1.5836}{75} + \frac{0.5739}{50}} 1.96 \right) = 1 - \alpha$$

$$p(0.1462 \leq (\mu_1 - \mu_2) \leq 0.8537) = 95\%$$

مثال تطبيقي 2:

اختيرت عينة 1 من الخراف التي تناولت الغذاء A ، وعينة أخرى 2 من الخراف تناولت الغذاء B..

إذا علمت أننا قمنا بقياس متوسط أوزان الخراف بعد فترة، وتحصلنا على المعلومات التالية:

A

B

$$\bar{x}_1 = 25.62$$

$$\bar{x}_2 = 28.33$$

$$\sum (X_{i1} - \bar{X}_1)^2 = 439.23$$

$$\sum (X_{i2} - \bar{X}_2)^2 = 1567.1096$$

$$n_1=13$$

$$n_2=15$$

- اوجد تقدير بفترة ثقة للفرق بين متوسطي المجتمعين. بمستوى ثقة 95%.

الحل:

$$\left\{ \begin{array}{l} \text{تباين المجتمعين مجهولين} \\ n_1, n_2 < 30 \end{array} \right.$$

فإننا نستخدم مقدرة تباين الفرق بين متوسطي مجتمعين

$$\sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}}$$

بدلاً من:

$$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

بما أن $n_1, n_2 < 30$ فإن توزيع الفرق يتبع توزيع ستيودنت: حيث:

$$p \left((\bar{x}_1 - \bar{x}_2) - \sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}} t_{1-\frac{\alpha}{2}} \leq (\mu_1 - \mu_2) \leq (\bar{x}_1 - \bar{x}_2) + \sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}} t_{1-\frac{\alpha}{2}} \right) = 1 - \alpha$$

حساب $t_{1-\frac{\alpha}{2}}$:

$$1 - \alpha = 0.95 \rightarrow \alpha = 0.05$$
 لدينا

ومنه:

$$t_{0.975}(v=n_1+n_2-2) = t_{0.975}(v = 26) = 2.056$$

حساب $\hat{S}_1^2$ و $\hat{S}_2^2$:

$$\begin{aligned}\hat{S}_1^2 &= \frac{\sum(X_{i1} - \bar{X}_1)^2}{n_1 - 1} \\ &= \frac{439.23}{12} = 36.6025\end{aligned}$$

$$\begin{aligned}\hat{S}_2^2 &= \frac{\sum(X_{i2} - \bar{X}_2)^2}{n_2 - 1} \\ &= \frac{1567.1096}{14} = 111.9364\end{aligned}$$

$$\begin{aligned}p\left((25.62 - 28.33) - \sqrt{\frac{36.6025}{13} + \frac{111.9364}{15}} 2.056 \leq (\mu_1 - \mu_2)\right. \\ \left.\leq (25.62 - 28.33) + \sqrt{\frac{36.6025}{13} + \frac{111.9364}{15}} 2.056\right) = 95\%\end{aligned}$$

$$p(-2.71 - (6.59) \leq (\mu_1 - \mu_2) \leq -2.71 + 6.59) = 95\%$$

$$p(-9.3 \leq (\mu_1 - \mu_2) \leq 3.88) = 95\%$$

طريقة أخرى للحل: نبحث مباشرة عن S^* مقدرة الانحراف المعياري للفرق بين متوسطي مجتمعين

$$\begin{aligned}p\left((\bar{x}_1 - \bar{x}_2) - S^* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{1-\frac{\alpha}{2}} \leq (\mu_1 - \mu_2)\right. \\ \left.\leq (\bar{x}_1 - \bar{x}_2) + S^* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{1-\frac{\alpha}{2}}\right) = 1 - \alpha\end{aligned}$$

$$S^* = \sqrt{\frac{\sum(x_{i1} - \bar{x}_1)^2 + \sum(x_{i2} - \bar{x}_2)^2}{n_1 + n_2 - 2}}$$

$$S^* = \sqrt{\frac{\sum 439.23 + 1567.1096}{13 + 15 - 2}}$$

$$S^* = \sqrt{\frac{\sum 439.23 + 1567.1096}{13 + 15 - 2}}$$

$$S^* = 8.784$$

$$p \left((-2.71) - 8.784 \sqrt{\frac{1}{13} + \frac{1}{15}} (2.056) \leq (\mu_1 - \mu_2) \right. \\ \left. \leq (-2.71) - 8.784 \sqrt{\frac{1}{13} + \frac{1}{15}} (2.056) \right) = 95\%$$

2- مجال الثقة للفرق بين نسبتي مجتمعين:

نُستخدم حدود الثقة للفرق بين نسبتي مجتمعين $p_1 - p_2$ لتقدير الفرق الحقيقي بين نسبتي مجتمعين مختلفين بناءً على بيانات عينتين عشوائيتين مستقلتين، تُحسب هذه الحدود عادةً باستخدام التقريب الطبيعي (للعينات الكبيرة) وفقاً للصيغة الرياضية التالية:

$$(p_1 - p_2) \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

ويكون مجال التقدير على النحو التالي:

$$\Pr \left[(p_1 - p_2) - Z_{\frac{\alpha}{2}} \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \leq \pi_1 - \pi_2 \leq (p_1 - p_2) + Z_{\frac{\alpha}{2}} \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \right] = 1 - \alpha$$

مثال تطبيقي:

عينة مكونة من 250 رجل، وعينة مكونة من 200 امرأة ، استخدموا تطبيق بريدي موب، فتبين أن 190 من الرجال و140 من النساء يفضلون التطبيق، قدر عند مستوى ثقة 95% الفرق بين نسبة الرجال ونسبة السيدات الذين يستخدمون بريدي موب ويفضلونه.

التقييم الذاتي للدرس التاسع

أولاً: أسئلة للفهم

- ماذا يعبر الفرق $\mu_1 - \mu_2$ ؟
- متى نستخدم توزيع t بدلاً من التوزيع الطبيعي (Z) عند تقدير الفرق بين متوسطين؟
- ما الشرط الأساسي لاستخدام التقريب الطبيعي لتقدير الفرق بين نسبتيين؟
- ماذا تعني "درجات الحرية" في حالة مقارنة متوسطي عينتين صغيرتين؟

ثانياً: أسئلة تحليلية

- لماذا في حالة $n_1 < 30$ و $n_2 < 30$ لا نستطيع استخدام التوزيع الطبيعي حتى لو عرفنا التباينين؟
- إذا كان فاصل الثقة للفرق بين متوسطين يحتوي على الصفر، ماذا يمكن أن نستنتج؟
- ماذا يحدث لعرض فاصل الثقة للفرق بين نسبتيين إذا زاد حجم إحدى العينتين مع بقاء النسبة ثابتة؟

ثالثاً: تطبيقات

التمرين 1:

لديك عينة من 100 رجل و 120 امرأة، تريد تقدير الفرق بين نسبة الرجال الذين يقرؤون الصحف يومياً ونسبة النساء الذين يقرؤون يومياً. توجد في العينة 55 رجلاً و 70 امرأة يفعلون ذلك. قدر مجال الثقة عند ثقة 95% للفرق بين المتوسطين.

التمرين 2:

في تجربة زراعية، استخدمت أرضان مختلفتان، سحبت عينة من 12 حبة قمح من الأولى و 15 حبة من الثانية، متوسط الإنتاج في العينة الأولى 50 كغ، وفي الثانية 45 كغ، والتباينان غير معروفين، والعينات صغيرة.

قم بتقدير مجال ثقة للفرق بين المتوسطين عند 99%.

رابعاً: اختيار الإجابة الصحيحة

عند بناء فاصل ثقة للفرق بين متوسطين بتباينين معلومين، نستخدم:

(أ) توزيع t

(ب) توزيع F

(ج) التوزيع الطبيعي Z

(د) توزيع مربع كاي

إذا كان فاصل الثقة 95% للفرق بين متوسطين هو $[-0.5, 2.5]$ ، فإن ذلك يعني:

- أ) الفرق يساوي صفر حتماً
- ب) لا توجد فروق معنوية عند مستوى 0.05 (لأن الصفر داخل الفترة).
- ج) الفرق أقل من -0.5
- د) الفرق أكبر من 2.5
- شرط استخدام التقريب الطبيعي للفرق بين نسبتيين يشمل:
- أ) كلا العينتين أكبر من 30
- ب) عدد النجاحات والإخفاقات في كل عينة ≥ 5
- ج) المجتمعان طبيعيان
- د) التباينان معلومان
- إذا زاد مستوى الثقة من 90% إلى 99%، فإن فاصل الثقة للفرق بين نسبتيين يصبح:
- أ) أضيق
- ب) أوسع
- ج) لا يتغير
- د) يختفي

الفصل الثالث: اختبار الفرضيات الاحصائية

مقدمة الفصل:

بعد أن تعرفنا في الدروس السابقة على التقدير النقطي والتقدير بفواصل الثقة كأدوات لاستنتاج معالم المجتمعات، ننتقل الآن إلى أداة إحصائية أخرى لا تقل أهمية: اختبار الفرضيات (Hypothesis Testing). بينما يهدف التقدير إلى إعطاء قيمة أو فترة للمعلمة، يهدف اختبار الفرضيات إلى اتخاذ قرار (قبول أو رفض) بشأن ادعاء أو فرضية حول معلمة المجتمع، بناءً على بيانات العينة.

ترجع الأسس النظرية الحديثة لاختبار الفرضيات إلى أعمال جيرزي نايمان (Jerzy Neyman) و إيفون بيرسون (Egon Pearson) في أوائل القرن العشرين، والتي طورت منهجية صارمة للتحكم في أخطاء القرار.

سنتناول في الدروس الأربعة التالية:

- **الدرس العاشر :** نظرية اختبار الفرضيات الأساسية، ومبادئ نايمان-بيرسون.
- **الدرس الحادي عشر :** اختبار الفرضيات لعينة واحدة (المتوسط، النسبة).
- **الدرس الثاني عشر :** اختبار الفرضيات لمقارنة عينتين (متوسطين، نسبتي).
- **الدرس الثالث عشر :** اختبار الفرضيات للتباين في مجتمع واحد ومقارنة تباينين.

هذه الدروس تشكل العمود الفقري للإحصاء الاستدلالي التطبيقي، وتستخدم في كل المجالات: الطب، الهندسة، الاقتصاد، العلوم الاجتماعية، ومراقبة الجودة.

الفصل الثالث: اختبار الفرضيات الإحصائية

الدرس العاشر: اختبار الفرضيات الإحصائية مع الإشارة الى نظرية نايمن بيرسون

المقدمة التعريفية

في هذا الدرس، سنُعرِّف مفهوم الفرضية الإحصائية، الفرضية الصفرية والبديلة، وأخطاء النوع الأول والنوع الثاني. سنُحلل منهجية نايمن-بيرسون في اختبار الفرضيات، بما في ذلك مفهوم منطقة الرفض، مستوى الدلالة، وقوة الاختبار. سنُطبق خطوات اختبار الفرضيات (صياغة الفرضيات، تحديد إحصائية الاختبار، تحديد منطقة الرفض، حساب القيمة، اتخاذ القرار). سنُقيّم العلاقة بين حجم العينة، مستوى الدلالة، وقوة الاختبار. سنُصمم أمثلة تطبيقية لاختبارات بسيطة.

الهدف العام للدرس

أن يُتقن الطالب بنهاية هذا الدرس المبادئ الأساسية لاختبار الفرضيات وفق منهجية نايمن-بيرسون، بحيث يتمكن من صياغة فرضيات صفرية وبديلة، وتحديد أخطاء النوع الأول والثاني، وحساب قوة الاختبار، واتخاذ القرارات الإحصائية السليمة في سياقات متنوعة.

الأهداف التفصيلية

- ❖ يُعرِّف الفرضية الصفرية H_0 والفرضية البديلة H_1 ، ويميز بين الاختبارات أحادية وثنائية.
- ❖ يُحدد أخطاء النوع الأول (α) والنوع الثاني (β)، ويشرح علاقتهما بمستوى الدلالة.
- ❖ يُطبق خطوات اختبار الفرضيات: صياغة، تحديد إحصائية الاختبار، منطقة الرفض.
- ❖ يُحلل تأثير حجم العينة ومستوى الدلالة على قوة الاختبار.
- ❖ يُقيّم متى يكون اختبار واحد أفضل من آخر باستخدام دالة القوة.

المحاور الأساسية للدرس

- ❖ مفهوم الفرضية الإحصائية
- ❖ أخطاء القرار
- ❖ منهجية نايمن-بيرسون
- ❖ خطوات اختبار الفرضيات العامة
- ❖ اختبارات معنوية أحادية وثنائية الجانب
- ❖ العلاقة بين فترات الثقة واختبارات الفرضيات

الوقت المتوقع: 4.5 ساعات بين 3 ساعات محاضرات و 1 سا و 30 حصة أعمال موجهة

1- مفهوم الفرضية الإحصائية

الفرضية الإحصائية هي عبارة أو ادعاء حول معلمة مجتمع (أو عدة معالم). في اختبار الفرضيات، نختبر عادة فرضيتين متضادتين:

- **الفرضية الصفرية : (H_0)** وهي الفرضية التي نبدأ بافتراض صحتها (غالباً تنص على عدم وجود تأثير أو عدم وجود فرق، أو تساوي معلمة لقيمة محددة (Prem & Christopher, 2010, p. 382).

(مثال: $H_0: \mu = \mu_0 = 20$)

- **الفرضية البديلة (H_1)** وهي الفرضية التي نعتبرها إذا رفضنا الفرضية الصفرية.
مثال:

$H_1: \mu \neq 20$ ثنائي الجانب

أو $H_1: \mu < 20$ أحادي الجانب اليساري

أو $H_1: \mu > 20$ أحادي الجانب اليميني

أمثلة:

عند اختبار متوسط عمر المصابيح: 10000 ساعة، والشركة تدي أن هذه المدة أطول من 10000 ساعة.

الفرضيات تكون بالشكل التالي:

$$\begin{cases} H_0: \mu = \mu_0 = 10000 \\ H_1: \mu > \mu_0 > 10000 \end{cases}$$

2- أخطاء القرار في اختبار الفرضيات

عند اتخاذ قرار بقبول أو رفض H_0 ، يمكن أن نرتكب نوعين من الأخطاء:

الجدول رقم 3: أخطاء القرار في اختبار الفرضيات.

القرار / الحقيقة	H_0 صحيحة	H_0 خاطئة
رفض H_0	خطأ النوع الأول (Type I Error)	قرار صحيح (رفض صحيح)
قبول H_0	قرار صحيح (قبول صحيح)	خطأ النوع الثاني (Type II Error)

- خطأ النوع الأول (α) رفض فرضية صحيحة، احتماله يُرمز له بـ α ، ويسمى مستوى الدلالة (Significance Level) يحدده الباحث مسبقاً (غالباً 0.05، 0.01، 0.10).
- خطأ النوع الثاني (β) قبول فرضية خاطئة، احتماله β قوة الاختبار (Power) هي $1-\beta$ ، أي احتمال رفض H_0 عندما تكون خاطئة (أي اكتشاف تأثير حقيقي). (Prem & Christopher, 2010, p. 386)
- ملاحظة: لا يمكن التحكم في كلا الخطأين في نفس الوقت لنفس العينة؛ تقليل α يؤدي عادة إلى زيادة β والعكس، الطريقة الوحيدة لتقليل كليهما هي زيادة حجم العينة n

3- منهجية نايمن-بيرسون (Neyman-Pearson Framework)

قدم نايمن وبيرسون (1928, 1933) إطاراً صارماً لاختبار الفرضيات:

1. تحديد (H_0) و (H_1)
 2. تحديد مستوى الدلالة α أقصى احتمال لخطأ النوع الأول نسمح به.
 3. اختيار إحصائية اختبار مناسبة مثل Z , t , F , χ^2
 4. تحديد منطقة الرفض (نطاق قيم إحصائية الاختبار التي تؤدي إلى رفض (H_0) بحيث يكون احتمال رفض (H_0) عندما تكون صحيحة يساوي α .
 5. حساب قيمة إحصائية الاختبار من العينة.
 6. اتخاذ القرار: إذا وقعت في منطقة الرفض نرفض (H_0)، وإلا نقبلها.
- مبرهنة نايمن-بيرسون الأساسية (للفرضيات البسيطة):
لنفرض أن لدينا فرضيتين بسيطتين:

$$H_0: \theta = \theta_0$$

$$H_1: \theta = \theta_1$$

فإن الاختبار الأقوى (ذو أعلى قوة) عند مستوى α هو الذي يرفض H_0 عندما:

$$\frac{L(\theta_0)}{L(\theta_1)} \leq k$$

حيث $L(\theta)$ هي دالة الاحتمال (likelihood)، و k ثابت يحدد بحيث يكون مستوى الدلالة α هذا هو اختبار نسبة الاحتمالات (Likelihood Ratio Test) للفرضيات البسيطة. (TAPER, Mark

L., PONCIANO, José M., et TOQUENAGA, Y, 2022, p. 62)

بالنسبة للفرضيات المركبة، يتم تعميم المبدأ باستخدام نسبة الاحتمالات المعممة (Generalized Likelihood Ratio).

دالة القوة (Power Function)

دالة القوة $\pi(\theta)$ تعطي احتمال رفض H_0 كدالة في المعلمة الحقيقية θ نحتاج أن تكون $\pi(\theta)$ منخفضة عندما θ تحت H_0 قريبة من α ومرتفعة عندما θ تحت H_1 قوة عالية.

4-خطوات اختبار الفرضيات

الخطوة 1 : صياغة الفرضيات (H_0) و (H_1)

الخطوة 2 : اختيار مستوى الدلالة α مثلاً 0.05

الخطوة 3 : تحديد إحصائية الاختبار المناسبة وتوزيعها تحت H_0 أمثلة:

- اختبار Z لمتوسط مع σ معلوم.
- اختبار t لمتوسط مع σ مجهول.

الخطوة 4 : تحديد منطقة الرفض (القيم الحرجة) من الجداول .

الخطوة 5 : حساب قيمة إحصائية الاختبار من بيانات العينة.

الخطوة 6 : اتخاذ القرار: إذا كانت قيمة الاختبار في منطقة الرفض، نرفض (H_0) ؛ وإلا نقبل (H_0)

الخطوة 7 : تفسير النتيجة في سياق المسألة. (Hansen, 2022, pp. 275-276)

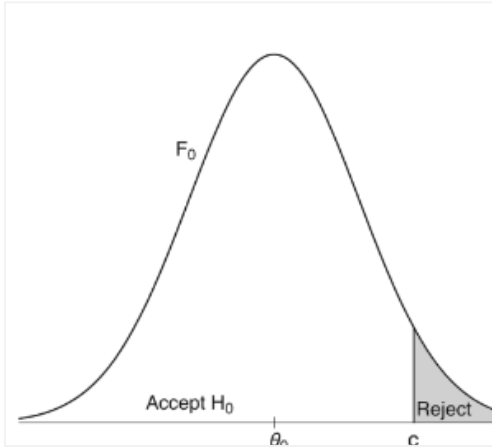
5-العلاقة بين فترات الثقة واختبارات الفرضيات

لفرضية ثنائية الجانب:

$$H_0: \theta = \theta_0$$

$$H_1: \theta \neq \theta_1$$

عند مستوى دلالة α ، فإن رفض H_0 يكافئ عدم وقوع θ_0 في مجال الثقة $(1-\alpha)$ للمعلمة θ .



التقييم الذاتي للدرس العاشر

أولاً: أسئلة للفهم

- ما الفرق بين الفرضية الصفريّة والفرضية البديلة؟
- عرّف خطأ النوع الأول وخطأ النوع الثاني مع مثال.
- ما المقصود بـ "قوة الاختبار"؟
- لماذا لا يمكن تقليل كلا الخطأين معاً دون زيادة حجم العينة؟
- ما هي العلاقة بين فاصل الثقة واختبار الفرضية الثنائية الجانب؟

ثانياً: اختيار الإجابة الصحيحة

مستوى الدلالة α هو احتمال:

- (أ) خطأ النوع الأول
 - (ب) خطأ النوع الثاني
 - (ج) قبول H_0 الصحيحة
 - (د) رفض H_0 الخاطئة
- إذا أردت اختبار فرضية أن متوسط المجتمع يساوي 100 مقابل أنه لا يساوي 100، فهذا اختبار:

- (أ) أحادي الجانب يميني
- (ب) أحادي الجانب يساري
- (ج) ثنائي الجانب
- (د) غير محدد

اختبار نايمن-بيرسون للفرضيات البسيطة يعتمد على:

- (أ) نسبة الاحتمالات
- (ب) مجموع المشاهدات
- (ج) التباين المشترك
- (د) الوسيط

الفصل الثالث: اختبار الفرضيات الإحصائية

الدرس الحادي عشر: الاختبارات الإحصائية المعلمية – اختبار الفرضيات للمتوسط والنسبة في مجتمع واحد

المقدمة التعريفية

بعد إتقان أساسيات اختبار الفرضيات، نطبقها الآن على حالات عملية: اختبار متوسط مجتمع واختبار نسبة مجتمع، سنميز بين اختبار المتوسط عندما يكون التباين معلوماً أو مجهولاً، سنطبق اختبار Z واختبار t للمتوسط، ونحلل شروط استخدام كل اختبار، سنقيم نتائج الاختبارات باستخدام القيم الحرجة والقيمة الاحتمالية (P -value). سنصمم خطوات كاملة لاختبار فرضيات حول نسبة النجاح في مجتمع برنولي.

الهدف العام للدرس

أن يتقن الطالب إجراء اختبارات الفرضيات للمتوسط (حالة تباين معلوم وحالة تباين مجهول) وللنسبة في مجتمع واحد، باستخدام التوزيعات المناسبة، مع القدرة على حساب القيمة الاختبارية وتفسير النتائج واتخاذ القرارات الإحصائية.

الأهداف التفصيلية

- ❖ يحدد متى يستخدم اختبار Z ومتى يستخدم اختبار t للمتوسط.
- ❖ يطبق اختبار Z للمتوسط (تباين معلوم) على بيانات حقيقية.
- ❖ يجري اختبار الفرضيات لنسبة المجتمع p باستخدام التقريب الطبيعي.
- ❖ يفسر قيمة (P -value) P ويقارنها مع α لاتخاذ القرار.

المحاور الأساسية للدرس

- ❖ اختبار الفرضيات لمتوسط مجتمع (σ معلوم)
- ❖ اختبار الفرضيات لمتوسط مجتمع (σ مجهول، $n \geq 30$ أو $n < 30$)
- ❖ اختبار الفرضيات لنسبة مجتمع (p)
- ❖ القيمة الاحتمالية (P -value) وأهميتها.

الوقت المتوقع: 4.5 ساعات بين 3 ساعات محاضرات و 1 ساعة و 30 دقيقة أعمال موجهة

1. اختبار المتوسط – تباين المجتمع معلوم (Prem & Christopher, 2010, p. 386)

الفرضيات:

• ثنائي الجانب:

$$H_0: \theta = \theta_0$$

$$H_1: \theta \neq \theta_1$$

• أحادي الجانب يميني:

$$H_0: \theta = \theta_0$$

$$H_1: \theta \geq \theta_1$$

• أحادي الجانب يساري:

$$H_0: \theta = \theta_0$$

$$H_1: \theta \leq \theta_1$$

إحصائية الاختبار:

$$Z_c = \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}} \sim N(0,1)$$

منطقة الرفض لمستوى دلالة α ، نرفض H_0 ونقبل H_1 إذا كان:

• ثنائي الجانب:

$$|Z_c| > Z_{1-\alpha/2}$$

• يميني:

$$Z_c > Z_{1-\alpha}$$

• يساري :

$$Z_c < -Z_{1-\alpha}$$

مثال تطبيقي:

مصنع ينتج قضبان حديد طولها المعلن 30 سم، انحراف معياري معروف 0.5 سم، عينة 40 قضبان أعطت متوسط 29.8 سم. اختبر عند $\alpha=0.05$ إذا كان الطول أقل من المعلن.

الحل:

$$\begin{cases} H_0: \mu = \mu_0 = 30 \\ H_1: \mu < \mu_0 < 30 \end{cases} \text{ 1/وضع الفرضيات:}$$

$$Z \text{ نحسب } n > 30 \begin{cases} \bar{x} = 29.8 \\ n = 40 \\ \sigma = 0.5 \\ \mu_0 = 30 \end{cases} \text{ لدينا:}$$

2/ حساب Z

$$Z_c = \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.5}{\sqrt{40}} = 0.07905$$

$$Z_c = \frac{29.8 - 30}{0.07905} = -2.53$$

3/ استخراج Z_{table}

$$Z_{table} = Z_{1-\alpha} = Z_{0.950} = 1.64$$

4/ اختبار الفرضيات:

$$Z_c = -2.53 < -Z_{table} = -1.64$$

نرفض H_0 ونقبل H_1 ، أي الطول أقل من المعلن.

2- اختبار المتوسط – تباين المجتمع مجهول (Prem & Christopher, 2010, p. 405)

نستخدم توزيع t بدلاً من Z

إحصائية الاختبار:

$$t_c = \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}} \sim t(n - 1)$$

مثال:

عينة 25 بطارية أعطت متوسط عمر 490 ساعة، انحراف معياري للعينة 12 ساعة.
- اختبر فرضية أن المتوسط 500 ساعة مقابل أنه أقل، عند $\alpha=0.05$

1/ وضع الفرضيات: $\begin{cases} H_0: \mu = \mu_0 = 500 \\ H_1: \mu < \mu_0 < 500 \end{cases}$

لدينا: $\begin{cases} \bar{x} = 490 \\ n = 25 \\ \sigma = 12 \\ \mu_0 = 500 \end{cases}$ $n < 30$ نحسب t

2/ حساب t

$$t_c = \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = 2.4$$

$$t_c = \frac{490 - 500}{2.4} = -4.16$$

3/ استخراج t_{table}

$$t_{table} = t_{1-\alpha}^{v=n-1=24} = t_{0.050}^{v=24} = -1.711$$

4/ اختبار الفرضيات:

$$t_c = -4.16 < t_{table} = -1.711$$

نرفض H_0 ونقبل H_1

3- اختبار النسبة (p)

اختبار أن نسبة النجاح في المجتمع تساوي P_0

إحصائية الاختبار:

إذا كان الاختبار ذو طرفين كما في الفرضية التالية:

$$\begin{cases} H_0: p = p_0 \\ H_1: p \neq p_0 \end{cases}$$

فعند مستوى دلالة α يكون قرار رفض أو عدم رفض الفرضية المبدئية H_0 وفقاً لقاعدة اتخاذ القرار التالية:

$$\begin{cases} Z = \frac{f - p_0}{\sqrt{v(f)}} \in [-Z_{1-\frac{\alpha}{2}}; +Z_{1-\frac{\alpha}{2}}] \rightarrow \overline{RH_0} \\ Z = \frac{f - p_0}{\sqrt{v(f)}} \notin [-Z_{1-\frac{\alpha}{2}}; +Z_{1-\frac{\alpha}{2}}] \rightarrow RH_0 \end{cases}$$

مع العلم أن:

$$v(f) = \frac{p_0(1 - p_0)}{n}$$

مثال:

ادعى مدير مدرسة أن نسبة النجاح في الامتحان 80%. عينة 200 طالب وجد 150 ناجحاً. اختبر عند $\alpha=0.05$ إذا كانت النسبة أقل من 80%.

4- القيمة الاحتمالية (P-value)

P-value هو احتمال الحصول على قيمة إحصائية اختبار متطرفة أو أكثر تطرفاً من القيمة المحسوبة،

بافتراض صحة H_0 . إذا كانت $P\text{-value} \leq \alpha$ نرفض H_0 .

في المثال الأخير، $P\text{-value} = P(Z < -1.768) \approx 0.0385 < 0.05$ ، مما يؤكد رفض H_0 .

التقييم الذاتي للدرس الحادي عشر

أسئلة للفهم

1. متى نستخدم اختبار t بدلاً من Z للمتوسط؟
2. ما شرط استخدام التقريب الطبيعي لاختبار النسبة؟
3. كيف نحسب القيمة P لاختبار ثنائي الجانب؟

تطبيق

تمرين 1 :

عينة من 36 شخصاً، متوسط ضغط الدم 12.5، انحراف معياري معروف 1.2 -
 اختبار الفرضية أن المتوسط 12 مقابل أنه يختلف، عند مستوى 0.01.

تمرين 2 :

عينة من 400 ناخب، 220 يؤيدون مرشحاً. اختبار فرضية أن النسبة 0.5 مقابل أنها أكبر من 0.5 -
 ما هو التوزيع المستخدم؟

تمرين 3:

إذا كان متوسط ربح سهم إحدى الشركات هو ($\mu_0 = 15$) في العام الماضي، تم أخذ عينة من $n=7$ مساهمين عن توقعاتهم متوسط ربح السهم في العام الحالي فوجد أنه ($\bar{x} = 17$) بانحراف معياري ($\sigma = 2$).
 هل توافق المساهمين في تقديرهم لارتفاع ربح السهم لهذا العام؟ (عند مستوى معنوية 5%).

تمرين 4:

1/ لغرض تقدير الانحراف المعياري لأوزان الخراف في مزرعة ما، تم اختيار عينة عشوائية مكونة من 16 خروفاً، فوجد أن الانحراف المعياري للعينة (2.4 كغ).
 المطلوب:

تقدير الانحراف المعياري لأوزان الخراف في المزرعة ضمن حدود ثقة مقدارها 99%.

2/ ادعى مسؤول التنظيم في ملعب الشهيد زوقاري الطاهر بغليزان أن معدل فترة الانتظار للدخول إلى المدرجات لا تزيد عن 15 دقيقة، ($\mu_0 = 15$)، ويهدف اختبار هذا الادعاء تم أخذ عينة من 10 مشجعين وتبين أن معدل انتظارهم هو ($\bar{x} = 16.5$) بانحراف معياري ($\sigma_{\bar{x}} = 4$).

الفصل الثالث: اختبار الفرضيات الاحصائية

الدرس الثاني عشر: اختبار الفرضيات للمقارنة بين متوسطين ونسبتين في المجتمع

المقدمة التعريفية

في هذا الدرس، نوسع اختبار الفرضيات لمقارنة معالم مجتمعين مستقلين، سنُقارن بين متوسطي مجتمعين (حالات تباينان معلومان، تباينان مجهولان وعينات كبيرة، تباينان مجهولان وعينات صغيرة مع افتراض تساوي التباينين، (سنطبق اختبار Z واختبار t لعينتين مستقلتين. سنحلل الفرق بين نسبتين مجتمعين باستخدام التقريب الطبيعي، سنقيم مدى ملاءمة استخدام التباين المشترك (pooled variance) وننفذ الاختبارات خطوة بخطوة على أمثلة واقعية.

الهدف العام للدرس

أن يُتقن الطالب إجراء اختبارات الفرضيات لمقارنة متوسطي مجتمعين (حالات مختلفة) ونسبتين مجتمعين، باستخدام التوزيعات المناسبة، مع تفسير النتائج واتخاذ قرارات إحصائية سليمة.

الأهداف التفصيلية

- ❖ يُحدد الصيغة المناسبة لاختبار الفرق بين متوسطين بناءً على معلومية التباينين وحجم العينتين.
- ❖ يُطبق اختبار Z للفرق بين متوسطين (تباينان معلومان أو مجهولان مع عينات كبيرة).
- ❖ يُطبق اختبار t لعينتين مستقلتين مع افتراض تساوي التباينين (التباين المشترك).
- ❖ يُجري اختبار الفرق بين نسبتين مستقلتين باستخدام التقريب الطبيعي.

المحاور الأساسية للدرس

- ❖ اختبار الفرق بين متوسطي مجتمعين.
- ❖ اختبار الفرق بين نسبتين.

الوقت المتوقع: 4.5 ساعات بين 3 ساعات محاضرات و 1 ساعة و 30 دقيقة أعمال موجهة

1- اختبار الفرق بين متوسطي مجتمعين

نريد اختبار الفرضية الصفرية التالية:

$$H_0: \mu_1 - \mu_2 = \Delta_0$$

غالباً ما يكون : ($\Delta_0 = 0$)

إحصائية الاختبار:

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - \Delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

مثال تطبيقي:

أجر اختبار ذكاء على 50 طالباً من مدرسة (أ) و 50 من مدرسة (ب) . حيث:

$$\bar{x}_1 = 105$$

$$\bar{x}_2 = 100$$

$$\sigma_1 = 15$$

$$\sigma_2 = 15$$

$$\alpha = 0.05$$

$$Z = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$Z = \frac{(105 - 100)}{\sqrt{\frac{225}{50} + \frac{225}{50}}}$$

$$= \frac{5}{\sqrt{9}} = 1.667$$

القيم الحرجة (الجدولية):

استخراج Z_{table}

$$Z_{table} = Z_{1-\frac{\alpha}{2}} = Z_{0.975} = 1.96$$

اختبار الفرضيات:

بما أن:

$$Z_c = 1.667 < Z_{table} = 1.96$$

نقبل H_0 أي لا يوجد فرق معنوي.

2-اختبار الفرق بين نسبتي

نختبر الفرضية التالية :

$$H_0: p_1 = p_2$$

$$\text{أي } (p_1 - p_2 = 0)$$

نقدر النسبة المشتركة :

$$\hat{p} = (x_1 + x_2) / (n_1 + n_2)$$

إحصائية الاختبار:

$$Z = \frac{(\bar{p}_1 - \bar{p}_2)}{\sqrt{\bar{p}_1(1 - \bar{p}_1) + (\frac{1}{n_1} + \frac{1}{n_2})}}$$

التقييم الذاتي للدرس الثاني عشر

أسئلة للفهم

- متى نستخدم التباين المشترك (pooled variance) في اختبار الفرق بين متوسطين؟
- ما الفرق بين اختبار Z واختبار t في مقارنة متوسطين؟
- كيف نحسب النسبة المشتركة لاختبار الفرق بين نسبتيْن؟

اختيار متعدد

في اختبار الفرق بين متوسطين بتباينين مجهولين وعينات صغيرة، نفرض:

(أ) استقلال العينتين

(ب) تساوي التباينين

(ج) توزيع طبيعي للمجتمعين

(د) كل ما سبق

إذا أردنا اختبار الفرق بين نسبتيْن مستقلتين، إحصائية الاختبار تتبع تقريباً:

A. t

B. $N(0,1)$

C. χ^2

D. F

الفصل الثالث: اختبار الفرضيات الإحصائية

الدرس الثالث عشر: اختبار الفرضيات للتباين في المجتمع

المقدمة التعريفية

في هذا الدرس، سنُعرِّف مفهوم اختبار الفرضيات للتباين في مجتمع واحد ومقارنة تباين مجتمعين، سنُميّز بين استخدام توزيع مربع كاي (χ^2) لاختبار تباين مجتمع واحد، وتوزيع F (اختبار فيشر) لمقارنة تباين مجتمعين، سنُطبق خطوات اختبار F على بيانات حقيقية، بدءاً من حساب التباينين، ثم حساب إحصائية الاختبار، ومقارنتها بالقيمة الجدولية، وصولاً إلى قبول أو رفض الفرضية الصفرية، سنُحلل نتائج الاختبار وحالات تطبيقه في البحوث النفسية والتربوية والصناعية. سنُقيّم القرار الإحصائي بناءً على المقارنة بين القيمة المحسوبة والقيمة الجدولية، هذا الدرس يمكنكم من الإجابة على أسئلة مثل: هل تباين درجات الأولاد يختلف عن تباين درجات البنات؟ وهل تباين إنتاج آلة معينة مستقر أم لا؟

الهدف العام للدرس

أن يُتقن الطالب بنهاية هذا الدرس إجراء اختبارات الفرضيات للتباين في مجتمع واحد (باستخدام توزيع χ^2) ولمقارنة تباين مجتمعين مستقلين (باستخدام توزيع F)، وذلك من خلال حساب إحصائية الاختبار، واستخراج القيم الحرجة من الجداول، واتخاذ القرار الإحصائي المناسب (قبول أو رفض الفرضية الصفرية)، وتفسير النتائج في سياقات عملية متنوعة.

الأهداف التفصيلية

- ❖ يُعرِّف الطالب إحصائية الاختبار.
- ❖ يُميّز الطالب بين اختبار التباين لمجتمع واحد.
- ❖ يُطبق الطالب خطوات اختبار مقارنة تباين عينتين مستقلتين
- ❖ يُحلل الطلاب مثلاً تطبيقياً.

المحاور الأساسية للدرس

- ❖ اختبار الفرضيات للتباين في المجتمع.

الوقت المتوقع: 4.5 ساعات بين 3 ساعات محاضرات و 1 ساعة و 30 دقيقة أعمال موجهة

تمهيد:

بعد أن تعلمنا في الدروس السابقة كيفية اختبار الفرضيات للمتوسط والنسبة والفرق بين متوسطين ونسبتين، ننتقل اليوم إلى جانب آخر لا يقل أهمية: اختبار الفرضيات المتعلقة بتباين المجتمع. يُستخدم التباين لقياس مدى تشتت البيانات حول المتوسط، وفي كثير من التطبيقات نريد معرفة هل تباين مجتمع معين يساوي قيمة محددة (مثل مراقبة جودة الإنتاج)، أو هل هناك فرق حقيقي بين تباين مجتمعين مستقلين (مثل مقارنة تجانس أداء مجموعتين تجريبية وضابطة).

اختبار الفرضيات للتباين في المجتمع

يستخدم اختبار التباين لأختبار الدلالة الاحصائية للفرق بين تبايني عينتين مستقلتين، فعند استخدام الباحث المجموعتان التجريبية والضابطة مثلاً يحتاج الى اختبار الفرق بين تباينهما لمعرفة فيما اذا كانتا مسحوبتان من نفس المجتمع ام من مجتمعين مختلفين ويتم الاختبار وفق الصيغة الآتية:

$$F = \frac{S_1^2}{S_2^2}$$

حيث:

F الاختبار الفائي

S_1^2 التباين الكبير

S_2^2 التباين الصغير

ثم تقارن القيمة المحسوبة مع القيمة الفائية النظرية (الجدولية) المستخرجة عند درجتي حرية (n_1-1) للبسط و (n_2-1) للمقام. (الشيحة، 2025، الصفحات 196-198).

مثال: وزع اختباراً نفسياً الى عينة من 25 ولد و 18 بنت وقد وجد ان تباين درجات الاولاد 54.8 وتباين درجات البنات 76.44 اختبر الفرق بين التباينين عند مستوى 0.05

$$H_0 : \delta_1^2 = \delta_2^2$$

$$H_1 : \delta_1^2 \neq \delta_2^2$$

$$F = \frac{s_1^2}{s_2^2} = \frac{76.44}{54.8} = 1.394 \quad \text{القيمة المحسوبة}$$

القيمة النظرية عند مستوى 0.05 و درجتى حرية 17, 24 هي 2.00

$$1.394 < 2.00$$

القيمة المحسوبة اصغر من القيمة الجدولية

تقبل افرضية الصفرية

ليس هناك فرق بين تباين المجموعتين

مثال تطبيقي:

طبق اختباراً نفسياً على مجموعتين الاولى من الافراد الاولى من العاطلين عن العمل والثانية من المنظمين في اعمال جيدة وتم الحصول على البيانات التالية، هل هناك فرق بين تباين المجموعتين اخبر ذلك عند

مستوى 0.05

x	y	X^2	y^2
23	66	529	4356
64	42	4096	1764
75	75	5625	5625
22	64	484	4096
41	48	1681	2304
79	52	6241	2704
32	54	1024	2916
14	60	196	3600
350	65	19876	4225
	72		5184
			36774

$$H_0 : \sigma_1^2 = \sigma_2^2$$

$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

$$F = \frac{S_1^2}{S_2^2}$$

$$S_x^2 = \sqrt{\frac{n \sum x^2 - (\sum x)^2}{n(n-1)}}$$

$$= \frac{8 \times 19876 - (350)^2}{8(8-1)} = \frac{159008 - 122500}{56}$$

تباين المجموعة الاولى

$$= \frac{36508}{56} = 651.93$$

$$S_y^2 = \frac{10 \times 36774 - (598)^2}{10(10-1)} = \frac{367740 - 357604}{90}$$

تباين المجموعة الثانية

$$= \frac{10136}{90} = 112.62$$

القيمة المحسوبة

$$F = \frac{S_1^2}{S_2^2} = \frac{621.93}{112.62} = 5.522$$

القيمة الفائية الجدولية عند مستوى 0.05 ودرجتي حرية 7 للبسط 90 للمقامي هي 3.29

$$5.522 > 3.29$$

القيمة الفائية اكبر من القيمة الجدولية

ترفض الفرضية الصفرية

الاستنتاج: هناك فرق دال احصائياً بين تبايني المجموعتين.

التقييم الذاتي للدرس الثالث عشر: اختبار الفرضيات للتباين في المجتمع

أولاً: أسئلة للفهم

- ما هو اختبار "ف" (F-test) المستخدم في هذا الدرس؟
- متى نضع التباين الأكبر في البسط عند حساب قيمة "ف"؟
- ماذا تعني درجات الحرية للبسط والمقام في اختبار "ف"؟
- كيف تقارن القيمة المحسوبة بالقيمة الجدولية لاتخاذ قرار قبول أو رفض الفرضية الصفرية؟
- إذا كانت القيمة المحسوبة لـ "ف" أقل من القيمة الجدولية، هل تقبل الفرضية الصفرية أم نرفضها؟

ثانياً: أسئلة تحليلية

- في المثال الأول (الأولاد والبنات)، كانت القيمة المحسوبة 1.394 والقيمة الجدولية 2.00. ماذا يستنتج الباحث؟ ولماذا؟
- في المثال الثاني (العاطلون عن العمل والمنظمون)، كانت القيمة المحسوبة 5.522 والقيمة الجدولية 3.29. ماذا يستنتج الباحث؟ وما معنى هذا الاستنتاج في واقع الدراسة؟
- لماذا نستخدم اختبار "ف" قبل إجراء اختبار "ت" (t-test) لمقارنة متوسطي مجموعتين؟ وما العلاقة بينهما؟
- كيف تتغير قيمة "ف" المحسوبة إذا زاد تباين إحدى المجموعتين بشكل كبير مع ثبات الأخرى؟
- ما أهمية افتراض أن العينتين مستقلتان في اختبار "ف"؟ وماذا لو كانت العينتان مرتبطتين؟

ثالثاً: تطبيقات**التمرين 1:**

لديك مجموعتان من الطلاب: المجموعة (أ) درجاتهم في اختبار الرياضيات، والمجموعة (ب) درجاتهم في نفس الاختبار. تريد معرفة هل هناك فرق حقيقي بين تباين درجات المجموعتين. صف الخطوات التي ستتبعها لإجراء اختبار "ف".

التمرين 2:

في مصنع لإنتاج المصابيح، تريد المقارنة بين تباين أعمار مصابيح خط إنتاج جديد مقابل خط إنتاج قديم. ما هي الفرضية الصفرية التي تضعها؟ وما هي الفرضية البديلة؟

التمرين 3:

بعد حساب قيمة "ف" المحسوبة، وجدتها تساوي 2.5، والقيمة الجدولية عند مستوى 0.05 تساوي 2.2. ماذا تقرر؟ وهل النتيجة دالة إحصائياً؟

الفصل الثالث: اختبار الفرضيات الاحصائية

الدرس الرابع عشر: بعض الاختبارات اللامعلمية (Test de Kolmogorov-Smirnov, Le test de Cramer -von Mises)

المقدمة التعريفية

في هذا الدرس، سنتناول اثنين من أبرز الاختبارات اللامعلمية، اختبار كولموجوروف-سميرنوف (Kolmogorov-Smirnov test) لمقارنة توزيع عينة بتوزيع نظري (KS) أحادي العينة) أو لمقارنة توزيع عينتين (KS ثنائي العينة)، واختبار كرامير-فون ميزيس (Cramér-von Mises test) مشابه لـ KS لكنه يعتمد على متوسط مربع الفروق بين التوزيعين.

الهدف العام للدرس

أن يُتقن الطالب استخدام اختباري كولموجوروف-سميرنوف وكرامير-فون ميزيس للتحقق من تطابق توزيع عينة مع توزيع نظري أو لمقارنة توزيع عينتين، وذلك من خلال حساب إحصائي الاختبار، ومقارنته بالقيم الحرجة، وتفسير النتائج.

الأهداف التفصيلية

- ❖ يُعرّف الطالب إحصائية الاختبار.
- ❖ يُميّز الطالب بين اختبارات المعلمية ولا المعلمية.
- ❖ يُطبّق الطالب خطوات اختبارات لامعلمية.

المحاور الأساسية للدرس

- ❖ اختبار كولموجوروف-سميرنوف.
- ❖ اختبار كرامير-فون ميزيس.

الوقت المتوقع: 4.5 ساعات بين 3 ساعات محاضرات و 1 ساعة و 30 دقيقة أعمال موجهة

تمهيد:

في الدروس السابقة، تعرفنا على الاختبارات المعلمية (Parametric tests) مثل اختبار t واختبار F واختبار Z ، والتي تعتمد على افتراضات معينة حول توزيع المجتمع (مثل افتراض التوزيع الطبيعي، تجانس التباينات، واستقلال المشاهدات). في العديد من التطبيقات العملية، قد لا تتوفر هذه الافتراضات، خاصة عندما:

- يكون حجم العينة صغيراً جداً بحيث لا يمكن التحقق من افتراض الطبيعية.
- تكون البيانات اسمية (Nominal) أو رتبية (Ordinal) وليست كمية.
- يكون توزيع المجتمع شديد الالتواء (Skewed) أو يحتوي على قيم متطرفة (Outliers).
- لا يمكن تحديد شكل التوزيع الاحتمالي للمجتمع.

1- الاختبارات اللامعلمية

هنا تأتي أهمية الاختبارات اللامعلمية (Nonparametric Tests)، والتي تُسمى أيضاً اختبارات التوزيع الحر (Distribution-free tests)، لأنها لا تعتمد على افتراضات محددة حول شكل توزيع المجتمع. تعتمد هذه الاختبارات بشكل أساسي على الرتب (Ranks) أو الإشارات (Signs) أو التكرارات، بدلاً من القيم الأصلية للبيانات.

مزايا الاختبارات اللامعلمية:

- لا تتطلب افتراض التوزيع الطبيعي.
- يمكن تطبيقها على البيانات الرتبية والاسمية.
- أكثر متانة (Robustness) ضد القيم المتطرفة.
- يمكن تطبيقها على العينات الصغيرة جداً (Sharma, 2010, p. 382).

عيوب الاختبارات اللامعلمية:

- أقل كفاءة (Efficiency) من الاختبارات المعلمية عندما تتحقق افتراضات الأخيرة (أي قد تفشل في رفض الفرضية الصفرية الخاطئة عندما يكون حجم العينة صغيراً).
- الكفاءة النسبية (Relative efficiency) للاختبارات اللامعلمية مقارنة بالمعلمية تكون غالباً أقل، لكنها تقترب من 1 عندما يكبر حجم العينة. (Sharma, 2010, p. 382).

2- اختبار كولموجوروف-سميرنوف أحادي العينة (One-sample KS test)

يُستخدم اختبار كولموجوروف-سميرنوف أحادي العينة للتحقق مما إذا كانت عينة $X_1, X_2, \dots, X_n$ تأتي من مجتمع يتبع توزيعاً نظرياً محدداً (عادةً مستمر بالكامل) مثل التوزيع الطبيعي، الأسّي، المنتظم، إلخ. تتمثل فرضيات الاختبار في:

الفرضية الصفرية H_0 : العينة تتبع التوزيع النظري المفترض $F(x)$.

الفرضية البديلة H_1 : العينة لا تتبع التوزيع النظري المفترض.

إحصائية الاختبار

نقوم أولاً بترتيب بيانات العينة تصاعدياً

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$$

نحسب دالة التوزيع التراكمي التجريبية – (Empirical Cumulative Distribution Function)

$$F_n(x) = \frac{\text{عدد المشاهدات } < x}{n}$$

دالة التوزيع النظري المفترض $F(x)$: مثل $\Phi\left(\frac{x-\mu}{\sigma}\right)$ للتوزيع الطبيعي

إحصائية كولموجوروف-سميرنوف (KS) أحادية العينة هي:

$$D_n = \sup_x |F_n(x) - F(x)|$$

$$D_n = \max \left\{ \max_{1 \leq i \leq n} \left| \frac{i}{n} - F(x_i) \right|, \max_{1 \leq i \leq n} \left| F(x_i) - \frac{i-1}{n} \right| \right\}$$

القيم الحرجة:

- توجد قيم جدولية لـ D_n عند مستويات دلالة α مختلفة (مثل 0.01، 0.05، 0.10).

- إذا كانت D_n المحسوبة أكبر من القيمة الجدولية، نرفض H_0

الجدول رقم 4: القيم الجدولية Kolmogorov-Smirnov table critical values

SAMPLE SIZE (N)	LEVEL OF SIGNIFICANCE FOR D = MAXIMUM [$F_0(X) - S_n(X)$]				
	.20	.15	.10	.05	.01
1	.900	.925	.950	.975	.995
2	.684	.726	.776	.842	.929
3	.565	.597	.642	.708	.828
4	.494	.525	.564	.624	.733
5	.446	.474	.510	.565	.669
6	.410	.436	.470	.521	.618
7	.381	.405	.438	.486	.577
8	.358	.381	.411	.457	.543
9	.339	.360	.388	.432	.514
10	.322	.342	.368	.410	.490
11	.307	.326	.352	.391	.468
12	.295	.313	.338	.375	.450
13	.284	.302	.325	.361	.433
14	.274	.292	.314	.349	.418
15	.266	.283	.304	.338	.404
16	.258	.274	.295	.328	.392
17	.250	.266	.286	.318	.381
18	.244	.259	.278	.309	.371
19	.237	.252	.272	.301	.363
20	.231	.246	.264	.294	.356
25	.210	.220	.240	.270	.320
30	.190	.200	.220	.240	.290

3- اختبار كولموغوروف-سميرنوف ثنائي العينة (Two-sample KS test)

يُستخدم لمقارنة ما إذا كانت العينتان $X_1, \dots, X_{n1}$ و $Y_1, \dots, Y_{n2}$ تأتيان من نفس التوزيع المستمر (وليس بالضرورة معرفة شكله).

يقوم الاختبار على الفرضيات التالية:

H_0 التوزيعان متماثلان أي $F_X(t) = F_Y(t)$ لكل t .

H_1 التوزيعان مختلفان.

إحصائية الاختبار

نحسب دالتين تجريبيتين: $F_{n1}(x)$ و $G_{n2}(x)$

$$D_{n1,n2} = \sup_x |F_{n1}(x) - G_{n2}(x)|$$

نخلط القيمتين ونرتبهما، ثم نحسب الفرق بين الدالتين التجريبيتين عند كل نقطة. أكبر قيمة مطلقة للفرق هي D ، القيمة الحرجة لثنائي العينة تقريباً عند: $(\alpha=0.05)$.

$$D_{crit} \approx 1.36 \sqrt{\frac{n1 + n2}{n1n2}}$$

4-اختبار كرامير-فون ميزيس (Cramér-von Mises)

اختبار كرامير-فون ميزيس هو بديل لاختبار KS ، حيث يستخدم متوسط مربع الفروق بين ECDF و CDF النظري بدلاً من أقصى فرق.

الجدول رقم: 5 القيم الحرجة Cramér-von Mises

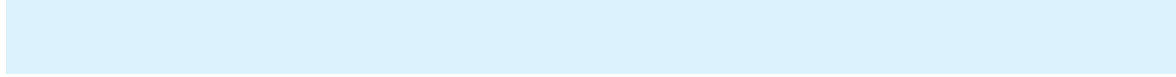
α F	.20	.15	.10	.05	.01
2	.138	.149	.162	.175	.186
3	.121	.135	.154	.184	.23
4	.121	.134	.155	.191	.28
5	.121	.137	.160	.199	.30
6	.123	.139	.162	.204	.31
7	.124	.140	.165	.208	.32
8	.124	.141	.165	.210	.32
9	.125	.142	.167	.212	.32
10	.125	.142	.167	.212	.32
11	.126	.143	.169	.214	.32
12	.126	.144	.169	.214	.32
13	.126	.144	.169	.214	.33
14	.126	.144	.169	.214	.33
15	.126	.144	.169	.215	.33
16	.127	.145	.171	.216	.33
17	.127	.145	.171	.217	.33
18	.127	.146	.171	.217	.33
19	.127	.146	.171	.217	.33
20	.128	.146	.172	.217	.33
30	.128	.146	.172	.218	.33
60	.128	.147	.173	.220	.33
100	.129	.147	.173	.220	.34

For $F > 100$ use values for $F = 100$; α = significance level.

إحصائية الاختبار أحادية العينة:

$$W^2 = N \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 dF(x)$$

$$W^2 = \frac{1}{12n} + \int_{i-1}^n F(xi) - \frac{2-1^2}{2n}$$



التقييم الذاتي للدرس الرابع عشر: اختبار كولموجوروف-سميرنوف واختبار كرامير-فون ميزيس

أولاً: أسئلة للفهم

- ما هو الهدف الرئيسي من اختبار كولموجوروف-سميرنوف أحادي العينة؟
- ماذا تقيس إحصائية "دي (D)" في اختبار KS ؟
- كيف يختلف اختبار كرامير-فون ميزيس عن اختبار KS من حيث المبدأ؟
- متى نستخدم اختبار KS ثنائي العينة بدلاً من أحادي العينة؟
- هل اختبار KS يتطلب معرفة التوزيع النظري مسبقاً؟ (نعم/لا)

ثانياً: أسئلة تحليلية

- إذا كانت القيمة المحسوبة لـ D أكبر من القيمة الجدولية، ماذا نستنتج؟
- ما ميزة اختبار كرامير-فون ميزيس مقارنة بـ KS ؟ وما عيبه؟
- لماذا نستخدم الاختبارات اللامعلمية مثل KS عندما لا نستطيع افتراض التوزيع الطبيعي؟
- في أي الحالات يكون اختبار KS غير مناسب لاختبار المطابقة؟

ثالثاً: تطبيقات

التمرين 1:

لديك عينة من 50 قياساً لارتفاع الأمواج في البحر، وتريد معرفة ما إذا كانت تتبع التوزيع الطبيعي. اتبع الخطوات لتطبيق اختبار KS مع ذكر ما تحتاجه من جداول.

التمرين 2:

باحث يريد مقارنة توزيع أعمار المرضى في مستشفيين مختلفين. ما الاختبار المناسب؟ وهل يحتاج إلى افتراض شكل التوزيع؟

رابعاً: اختيار الإجابة الصحيحة

إحصائية اختبار كولموجوروف-سميرنوف أحادي العينة تعتمد على:

- A. مجموع الفروق المربعة
- B. أقصى مسافة رأسية بين التوزيعين
- C. مجموع القيم المطلقة للفروق
- D. الفرق بين الوسيطين

عندما تكون القيمة المحسوبة D أصغر من القيمة الجدولية، فإننا:

A. نرفض الفرضية الصفرية

B. نقبل الفرضية الصفرية

C. نعيد الحساب

D. نرفض الفرضية البديلة

اختبار كرامير-فون ميزيس يختلف عن KS في أنه يستخدم:

A. أقصى مسافة

B. متوسط مربع المسافات

C. (ج) مجموع القيم المطلقة

D. (د) الفروق بين رتب العينات

الفصل الثالث: اختبار الفرضيات الاحصائية

الدرس الخامس عشر: اختبار (Test de Wilcoxon - Mann - Whitney) - (Test de Kruskal-Wallis) -

المقدمة التعريفية

أن يُتقن الطالب استخدام اختباري ويلكوكسون-مان-ويتني (لعينة مستقلة) وكروسكال-واليس (لثلاث عينات فأكثر) لمقارنة وسطيات مجتمعات عندما لا تتحقق افتراضات الاختبارات المعلمية.

الهدف العام للدرس

أن يُتقن الطالب استخدام اختباري كولموجوروف-سميرنوف وكرامير-فون ميزيس للتحقق من تطابق توزيع عينة مع توزيع نظري (اختبار المطابقة goodness-of-fit) أو لمقارنة توزيع عینتين، وذلك من خلال حساب إحصائي الاختبار، ومقارنته بالقيم الحرجة، وتفسير النتائج.

الأهداف التفصيلية

- ❖ يُعرّف الطالب متى يُستخدم اختبار ويلكوكسون-مان-ويتني بدلاً من t-test
- ❖ يحسب رتب العينات ويُطبق صيغة إحصائية U
- ❖ يُقارن بين اختبار ويلكوكسون-مان-ويتني واختبار t
- ❖ يُطبق اختبار كروسكال-واليس لمقارنة ثلاث مجموعات.
- ❖ يحسب إحصائية H ويستنتج القرار.

المحاور الأساسية للدرس

- ❖ اختبار ويلكوكسون-مان-ويتني (Wilcoxon-Mann-Whitney)
- ❖ حساب رتب العينات وإحصائية U
- ❖ اختبار كروسكال-واليس (Kruskal-Wallis)
- ❖ حساب إحصائية KW وتوزيعها التقريبي (χ^2)

الوقت المتوقع: 4.5 ساعات بين 3 ساعات محاضرات و 1 سا و 30 حصة أعمال موجهة

1-اختبار (WMW) (Test de Wilcoxon - Mann - Whitney)

يعرف أيضاً باختبار مجموع الرتب (Rank-sum test) يُستخدم لمقارنة عينتين مستقلتين لاختبار ما إذا كانت قيم إحدى العينتين تميل إلى أن تكون أكبر من الأخرى (أي اختبار الفرضية H_0 توزيعات العينتين متماثلة مقابل H_1 توزيعاتهما مختلفة – أو وسطياتهما مختلفة). هو البديل اللامعلمي الأكثر شيوعاً لاختبار t لعينتين مستقلتين.

يقوم الاختبار على الفرضيات التالية:

H_0 متوسطات الرتب متساوية (أي التوزيعان متماثلان).

H_1 أحد التوزيعين يميل إلى قيم أكبر.

نقوم بترتيب جميع قيم العينتين معاً تصاعدياً، ونعطي رتباً (في حالة التعادل، نأخذ متوسط الرتب). ثم نحسب مجموع الرتب لكل عينة على حدة. إذا كانت إحدى العينتين تميل إلى القيم الأكبر، فسيكون مجموع رتبها أكبر.

إحصائية U

لنفرض حجم العينة الأولى n_1 والثانية n_2 مجموع الرتب للعينة الأولى هو R_1 إحصائية مان-ويتني U تُحسب:

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

$$U_1 = R_1 - \frac{n_1(n_1 + 1)}{2}, \quad U_2 = R_2 - \frac{n_2(n_2 + 1)}{2}$$

قيمة U المستعملة هي $\min(U_1, U_2)$ (123-128 الصفحات، 2014، الكنانى)

القيم الحرجة

توجد جداول خاصة لـ U لكن للعينات الكبيرة ($n_1, n_2 \geq 20$) ، يمكن تقريب U بالتوزيع الطبيعي:

$$Z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

2- اختبار كروسكال-واليس (Kruskal-Wallis)

يعمم اختبار ويلكوكسون-مان-ويتني إلى $k \geq 3$ عينات مستقلة، يُستخدم لاختبار ما إذا كانت العينات تأتي من نفس التوزيع (أو لها نفس الوسيط)، هو البديل اللامعلمي لتحليل التباين أحادي الاتجاه (One-way ANOVA).

يقوم الاختبار على الفرضيات التالية:

H_0 جميع العينات تأتي من نفس التوزيع (وسيطات متساوية)

H_1 على الأقل عينتان تختلفان في التوزيع.

نقوم بترتيب جميع القيم من كل المجموعات معاً، ونعطي رتّباً، ثم نحسب مجموع الرتب لكل مجموعة إحصائية كروسكال-واليس.

$$KW = \frac{12}{T(T+1)} \left[\sum_{i=1}^P \frac{R_i^2}{n_i} \right] - 3(T+1) \rightarrow \chi^2_{(p-1)}$$

حيث: R_i^2 مربع مجموع رتب المشاهدات المقابلة للفصل i

n_i عدد المشاهدات المقابلة للفصل i

P تعبر عن الدورة = 4 في المشاهدات الفصلية، 12 في المشاهدات الشهرية، ... الخ.

KW يتبع توزيع كاي تربيع بدرجة حرية $P-1$ $\chi^2_{(p-1)}$ (الشريحة، 2025، الصفحات 237-240).

التقييم الذاتي للدرس الخامس عشر: اختبار ويلكوكسون-مان-ويتني واختبار كروسكال-واليس

أولاً: أسئلة للفهم

- ما هو البديل اللامعلمي الأكثر شيوعاً لاختبار t لعينتين مستقلتين؟
- على أي شيء يعتمد اختبار ويلكوكسون-مان-ويتني بدلاً من القيم الأصلية؟
- ماذا يمثل اختبار كروسكال-واليس مقارنة بتحليل التباين (ANOVA)؟
- متى نستخدم اختبار كروسكال-واليس بدلاً من ويلكوكسون-مان-ويتني؟
- ما معنى "الرتب (Ranks)" في سياق هذه الاختبارات؟

ثانياً: أسئلة تحليلية

- لماذا تعتبر الاختبارات المعتمدة على الرتب أقل تأثراً بالقيم المتطرفة من الاختبارات المعتمدة على المتوسطات؟
- في اختبار ويلكوكسون-مان-ويتني، إذا كان مجموع رتب إحدى العينتين أكبر بكثير من الأخرى، ماذا يستنتج الباحث؟
- ماذا تفعل عند وجود قيم متكررة (تعاذل) عند ترتيب الرتب؟
- كيف يرتبط اختبار كروسكال-واليس بتوزيع مربع كاي؟
- لماذا نفضل استخدام الاختبارات اللامعلمية عندما تكون العينات صغيرة جداً (مثل أقل من 10) ولا نستطيع التأكد من الطبيعية؟

ثالثاً: تطبيقات

التمرين 1:

لديك عينتان مستقلتان: الأولى من مرضى تناولوا دواءً جديداً، والثانية مجموعة ضابطة (دواء وهمي). قياسات الشفاء هي أرقام ترتيبية (سيء، متوسط، جيد، ممتاز). ما الاختبار المناسب لمقارنة المجموعتين؟ احسب الاختبار مع اتباع جميع الخطوات.

التمرين 2:

باحث يريد مقارنة فعالية ثلاثة أساليب تدريس مختلفة على أداء الطلاب، والبيانات لا تتبع التوزيع الطبيعي. ما الاختبار الذي سيستخدمه؟ وما التوزيع المرجعي لقيمة الاختبار؟

رابعاً: اختيار الإجابة الصحيحة

اختبار ويلكوسون-مان-ويتني يستخدم لـ:

أ) مقارنة تباين عينتين

ب) مقارنة متوسطي عينتين مرتبطتين

ج) مقارنة توزيع عينتين مستقلتين

د) اختبار مطابقة توزيع عينة مع نظري

إحصائية الاختبار في ويلكوسون-مان-ويتني تعتمد بشكل أساسي على:

أ) مجموع الرتب

ب) المتوسطات الحسابية

ج) الانحرافات المعيارية

د) القيم العظمى

اختبار كروسكال-واليس هو تعميم لـ:

أ) اختبار t لعينة واحدة

ب) اختبار ويلكوسون-مان-ويتني

ج) اختبار كولموجوروف-سميرنوف

د) اختبار كرامير-فون ميزيس

في اختبار كروسكال-واليس، إحصائية H تتبع تقريباً توزيع:

أ) طبيعي معياري

ب) مربع كاي

ج) ت

د) ف

أحد مزاي اختبار ويلكوكسون-مان-ويتني هو:

أ) أكثر قوة من اختبار t عندما تتحقق الطبيعية

ب) لا يتطلب افتراض طبيعة التوزيع

ج) يعمل فقط مع العينات الكبيرة

د) يعطي نتائج دقيقة دوماً

الخاتمة:

بإتمام هذه المطبوعة، يكون الطالب قد اكتسب مجموعة متكاملة من المفاهيم والأدوات الإحصائية اللازمة للتعامل مع مسائل الاستدلال الإحصائي، بدأنا بتوزيعات المعاينة المختلفة (للمتوسط، النسبة، الفروق، والتباين) التي تُمثل الأساس النظري لأي عملية استدلال. ثم انتقلنا إلى نظرية التقدير، حيث تعرفنا على خصائص المقدرات الجيدة (عدم التحيز، الاتساق، الكفاءة، الكفاية، التقارب)، وعلى طريقتين أساسيتين لاستنتاج المقدرات هما طريقة العزوم وطريقة المعقولية العظمى، بالإضافة إلى متباينة كرامير-راو التي تحدد الحد الأدنى لتباين أي مقدر غير متحيز، في الجانب التطبيقي، تعلمنا كيفية بناء فترات الثقة للمتوسط، النسبة، الفرق بين متوسطين ونسبتين، التباين، ونسبة التباينين. واختتمنا باختبارات الفرضيات وفق إطار نايمن-بيرسون، بما في ذلك اختبارات المتوسط والنسبة لعينة واحدة، ومقارنة متوسطين ونسبتين، واختبارات التباين (مربع كاي و F)، نأمل أن تشكل هذه المطبوعة مرجعاً مفيداً للطلاب والباحثين في التحليل الإحصائي، وأن تساهم في تطوير مهاراتهم في اتخاذ القرارات العلمية السليمة المستندة إلى بيانات حقيقية.

قائمة المراجع:

- Prem, S., & Christopher, J. (2010). *Introductory Statistics*. John Wiley & Sons.
- الحاج خليفة. (2019). *دروس وأعمال موجهة في الإحصاء 3*، مطبوعة موجهة إلى طلبة السنة الثانية والسنة الثالثة LMD علوم اقتصادية، علوم التسيير والعلوم التجارية، جامعة مستغانم، الجزائر.
- شريفية بوالشعور. (2018). *بوالشعور شريفية، مطبوعة دروس في الإحصاء 3*، موجهة لطلبة السنة الثانية علوم اقتصادية، مالية، تجارية، وعلوم التسيير، جامعة سكيكدة، الجزائر.
- محمد عبد المنعم ثروت. (2023). *مدخل حديث للإحصاء والاستدلال*، الطبعة 3، دار كعبيبان، الرياض، المملكة العربية السعودية.
- Casella, G., & Berger, R. (2024). *Statistical inference*. Chapman and Hall/CRC.
- GIRI, Narayan, C. (1993). *Introduction to Probability and Statistics: Revised and Expanded*. STATISTICS TEXTBOOKS AND MONOGRAPHS, . Taylor & Francis.
- Hansen, B. (2022). *Probability and Statistics for Economists*. Princeton University Press.
- LeBlanc, D. (2004). *Statistics: concepts and applications for science*. London, UK: Jones and Bartlett.
- Mátyás, L. (1999). *Generalized method of moments estimation (Vol. 5)*. Cambridge University Press.
- Montgomery, D., & Runger, G. (2019). *Applied statistics and probability for engineers*. John wiley & sons.
- Nielsen, F. (2013). *Cramér-Rao lower bound and information geometry*. In *Connected at Infinity II: A Selection of Mathematics by Indians (pp. 18-37)*. Gurgaon: Hindustan Book Agency.
- Sharma, J. (2010). *Business Statistics*. London, United Kingdom: Dorling Kindersley.

evidential. (2022) TAPER, Mark L., PONCIANO, José M., et TOQUENAGA, Y. *statistics, model identification, and science. Frontiers in Ecology and Evolution*, 2022, vol. 10, p. 883456.

أمين تمار . (2020). محاضرات وتمارين في الإحصاء 3، مطبوعة موجهة إلى طلبة السنة الثانية ليسانس علوم اقتصادية، علوم التسيير والعلوم التجارية، جامعة البليدة 2، الجزائر.
إياد محمد الهوبي. (2014). الإحصاء التطبيقي، الطبعة الأولى، الكلية الجامعية للعلوم والتكنولوجيا، خان يونس، فلسطين.

بلفضيل كمال. (2025). محاضرات في الإحصاء 4. مطبوعة بيداغوجية: جامعة غليزان.
جلال مصطفى الصياد، و جلال مصطفى مصطفى . (بلا تاريخ). مقدمة في طرق المعاينة الإحصائية، مكتبة الصباح، الطبعة الأولى.

جمعة عبدالسلام عبد السلام . (2026). دليل التحليل الإحصائي الميسر للعلوم التطبيقية. القاهرة، مصر: دار العربي للنشر والتوزيع.

جميل أحمد. (2010). أساليب المعاينة القياس وتحليل البيانات. دراسات العدد الاقتصادي، 1(1)، الصفحات 181-207.

حمزة بن سبع. (2020). دروس ومحاضرات في الإحصاء الاستدلالي (إحصاء 3) مع أمثلة وتمارين، مطبوعة بيداغوجية، جامعة الجزائر 3، الجزائر.

زهراوي عفاف. (2020). إحصاء 3، دروس على الخط: الدرس الثالث: التقديرات، جامعة قسنطينة، السداسي الأول.

عايد كريم عبدعون الكنائي. (2014). مقدمة في الإحصاء وتطبيقات SPSS. عمان، الأردن: دار اليازوردي العلمية للنشر والتوزيع.

عبد الحفيظ مصطفى. (2002). الاستدلال الإحصائي ج2. القاهرة، مصر: مجموعة النيل العربية طباعة نشر توزيع.

عبد الحفيظ محمد فوزي، و عبد الحفيظ مصطفى. (2000). الاستدلال الإحصائي. مصر: مجموعة النيل العربية طباعة نشر توزيع.

عبدالله بن عبدالكريم الشيحة. (2025). تصميم التجارب الإحصائية وتحليلها. الرياض، المملكة العربية السعودية: العبيكان للنشر.

فتحي عاروري. (2013). المعاينة الإحصائية طرقها واستخداماتها. عمان، الأردن: دار الأكاديميون للنشر والتوزيع.

- كالجيان هاري ، و أوتس ولاس . (2001). مقدمة في الاقتصاد القياسي: مبادئ وتطبيقات، ترجمة: المرسي السيد حجازي، عبد القادر محمد عطية، مراجعة: حمد بن سليمان البازعي، جامعة الملك سعود، المملكة العربية السعودية.
- مبارك اسبر ديب. (2009). مبادئ في الاحتمالات والإحصاء. سوريا: مديرية الكتب والمطبوعات بجامعة تشرين.
- محمد ابو صالح. (2020). مبادئ الاحصاء. الاردن: دار اليازوري.
- نبيل جمعة صالح النجار. (2015). الاحصاء التحليلي مع تطبيقات برمجية SPSS، دار ومكتبة الحامد، عمان، الأردن.
- يداوي محمد، دوة محمد. (2016). الإحصاء المطبق في العلوم الاجتماعية، دار اليازوري العلمية، عمان، الأردن.

