المحاضرة الثالثة:

توزيع المعاينة

ثانيا: توزيع المعاينة:

يُعد فهم توزيعات المعاينة مدخلاً مهماً لفهم موضوع الإستدلال الإحصائي (الإستنتاجي)، حيث يستخدم توزيع إحصاءات العينة بشكل واسع في موضوع اختبار الفرضيات الإحصائية.

ويعرف توزيع المعاينة بأنه عبارة عن: "توزيع كافة القيم المحتملة التي يمكن افتراضها بإحصاءة ما محسوبة على أساس عينات مختلفة بنفس الحجم اختيرت عشوائياً من نفس المجتمع".

وبناءاً على ما تقدم يتضح بأن توزيعات المعاينة قد يتم بناؤها تجريبياً عند سحب العينات من مجتمع متقطع ومحدود، ولبناء توزيع المعاينة نتبع الخطوات التالية:

أ- سحب كافة العينات بحجم (n) مع الإرجاع أو بدون إرجاع من مجتمع ومتقطع حجمه (N)؛ فلو أردنا سحب عينة من المجتمع حجمها (n) و كان حجم المجتمع (N)، وكان السحب مع الإرجاع فعدد الطرق الممكنة لسحب العينة هو (N^n)، أما لو كان السحب بدون إرجاع فعدد الطرق الممكنة هي: $C_n^N = \dfrac{N!}{(N-n)!n!}$ ؛ فمثلاً لو كان حجم المجتمع 10 و حجم العينة 3، و كان السحب مع الإرجاع فيمكن سحب $1000 = 10^3$ عينة مختلفة، أما إذا كان السحب بدون إرجاع فيمكن سحب $120 = C_3^{10}$ عينة مختلفة.

ب- حساب متوسط المجتمع (μ) وتباينه (σ^2)؛

ج- حساب الإحصاءة $(\overline{X})$ لكل عينة؛

د- تكوين جدول توزيع تكراري للقيم المختلقة للإحصاءة المحسوبة؛

هـ- حساب متوسط المتوسطات $(\mu_{\bar{x}})$ وتباينه $(\sigma_{\bar{x}}^2)$؛

وحيث أن التوزيع الطبيعي يكتسي أهمية واسعة في التطبيقات الاحصائية الواسعة، وسنذكر الآن بدون برهان نظرية هامة تسمى نظرية النهاية المركزية، والتي تلعب دوراً هاماً في الإحصاء التطبيقي، حيث تمكننا من تقريب أي توزيع احتمالي للتوزيع الطبيعي عند تحقق شروط معينة.

نظرية: ليكن $\{X_i\}_{i=1}^{n}$ مجموعة من المتغيرات المستقلة، و التي تتبع نفس التوزيع وكل منها بوسط (μ) وتباين(σ^2) إذا كان: $S_n = \sum X_i$ فإن: $\dfrac{S_n - n\mu}{\sigma\sqrt{n}} \sim N(0,1)$.

إن للتوزيع الطبيعي دور مهم آخر يتمثل في اشتقاق توزيعات احتمالية متعلقة بالمعاينة، و ذلك عند عينة صغيرة حجمها $(N < 30)$ تسحب عادة من مجتمع طبيعي وهذه التوزيعات هي:

- توزيع كاي مربع (χ^2)؛
- توزيع ستودنت (t)؛
- توزيع فيشر (F).

1- توزيع كاي مربع (χ^2) **Chi-Square Distribution**:

إن الأساس النظري لاشتقاق توزيع كاي مربع هو التوزيع الطبيعي، بمعنى آخر أن توزيع كاي مربع مشتق من الدرجة المعيارية (Z)، وعلى افتراض أنه إذا كان لدينا المتغير(X) كمتغير عشوائي يخضع للتوزيع الطبيعي بمتوسط حسابي (μ) وتباين (σ^2) أي أن: $X \sim N(\mu,\sigma^2)$:

ومنه تكون الدرجة المعيارية (Z) كالآتي:

$$Z = \frac{X-\mu}{\sigma} \sim N(0\,,1)$$

عليه تعد الدرجة المعيارية (Z) متغيراً عشوائياً يخضع للتوزيع الطبيعي المعياري (القياسي) بمتوسط حسابي يساوي (صفر) وتباين (1).

وبتربيع قيمة الدرجة المعيارية (Z) نحصل على المتغير (Z^2)، وعند البحث عن التوزيع الإحتمالي للمتغير العشوائي (Z^2) سيتضح بأنه يخضع لتوزيع كاي مربع (χ^2) بدرجة حرية واحدة، أي أن:

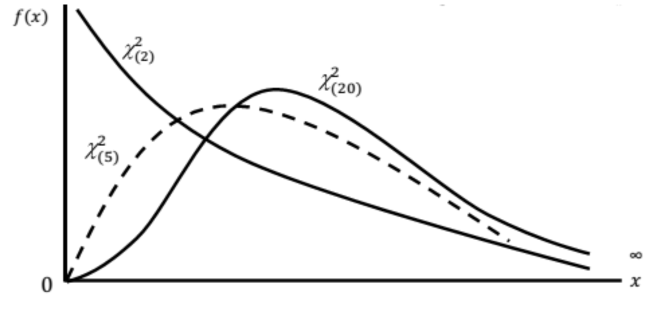
$$\because Z = \frac{X-\mu}{\sigma} \sim N(0,1)$$

$$\therefore Z^2 = \left(\frac{X-\mu}{\sigma}\right)^2 \sim \chi^2_{(1)}$$

وعند اختبار عينة عشوائية قوامها مشاهدتين ($n=2$) و هما (X_1, X_2) من بين مشاهدات المتغير العشوائي (X) الذي يخضع للتوزيع الطبيعي بمتوسط حسابي (μ) وتباين (σ^2).

1-1- خصائص توزيع كاي مربع: يتميز توزيع كاي مربع بالخصائص التالية:

- إن قيمة كاي مربع موجبة دائما، أي أن ($\chi^2 > 0$)؛
- إن الوسط الحسابي (μ) للتوزيع هو عدد درجات رية التوزيع (ν)؛
- إن تباين التوزيع (σ^2) هو ضعف عدد درجات الحرية (2ν)؛
- إن الإنحراف المعياري للتوزيع (σ) هو جذر التباين ($\sqrt{2\nu}$)؛
- إن منحنى توزيع كاي مربع (χ^2) ملتوي نحو اليمين (التواء موجب)؛
- يختلف شكل منحنى توزيع كاي مربع (χ^2) باختلاف درجات الحرية، والرسم الموالي يبين شكل التوزيع لدرجات حرية مختلفة:



ونعبر عن قيمة المتغير العشوائي الذي يتبع توزيع كاي مربع بالرمز $\chi^2_{(\alpha,\nu)}$ وهي القيمة التي يقع على يمينها مساحة (α) على منحنى كاي مربع بدرجة حرية (ν)، وتحسب قيم هذا المتغير العشوائي والإحتمالات له باستخدام الجدول الخاص به.

مثال: أوجد القيم التالية: $\chi^2_{(0.1\ ,\ 17)}$ ، $\chi^2_{(0.995\ ,\ 8)}$ ، $\chi^2_{(0.95\ ,\ 7)}$

الحل:

لإيجاد $\chi^2_{(0.95\ ,\ 7)}$ نختار من العمود الأيسر(عمود درجة الحرية) القيمة 7 فنجد على يمينه صف من الأعداد، ومن الصف العلوي (صف المساحات أي الإحتمالات) عن القيمة 0.95 فنجد اسفلها عمود من الأعداد، فتكـــون القيمة الموجود في تقاطع الصف والعمود هي القيمة المطلوبة، وبالتالي يكون:

$\chi^2_{(0.95\ ,\ 7)} = 2.167$ وكذلك: $\chi^2_{(0.995\ ,\ 8)} = 1.344$ ، $\chi^2_{(0.1\ ,\ 17)} = 24.769$.

1-2- استخدامات توزيع كاي مربع : إن لتوزيع كاي مربع استخدامات كثيرة منها:

- اختبار قيمة تباين مجتمع طبيعي التوزيع؛
- اختبار حسن المطابقة (توفيق البيانات)؛
- اختبار الإستقلالية؛
- اختبار تجانس عدة تقديرات مستقلة لتباين المجتمع الطبيعي؛
- اختبار معامل الإرتباط؛
- ايجاد فترات الثقة لتباين المجتمع.

2- توزيع ستودنت (t) t-Distribution:

إن الأساس لاشتقاق توزيع (t) هو في الواقع توزيع مشتق من حاصل قسمة متغيرين مستقلين أولهما يمثل المتغير العشوائي (Z) الذي يخضع للتوزيع الطبيعي المعياري بمتوسط حسابي يساوي (0) وتباين يساوي (1)، أمـــا المتغير العشوائي الثاني فهو في المقام يمثل الجدر التربيعي للمتغير (χ^2) الذي يخضع لتوزيع كاي-مربـــع بدرجـــة حرية (n) مقسوم على درجات حريته أي أن:

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

$$\chi^2 \sim \chi^2_n$$

يعد كل من المتغيرين العشوائيين (Z) و (χ^2_n) مستقلين عن بعضهما، عندئد يكون المتغير العشوائي (t) معرفاً بالصيغة التالية:

$$t = \frac{Z}{\sqrt{\chi^2_{(n)}}} \sim t_{(n)}$$

من الصيغة السابقة يتضح بأن المتغير العشوائي (t) يخضع إلى توزيع (t) بدرجة حرية مساوية إلى (n).

يستخدم توزيع (t) في حالة العينات الصغيرة وعلى وجه التحديدفي حالة العينات التي أقل من 30 مشاهدة $(n < 30)$.

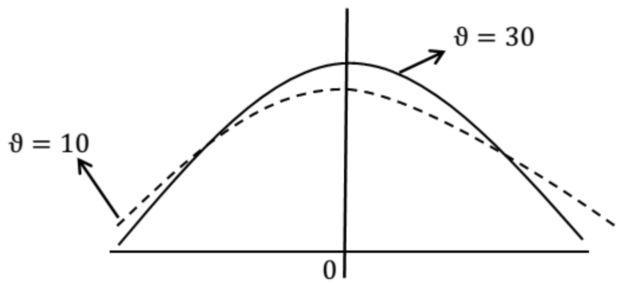
ويعود الفضل الأول للعالم (Gosset) في اكتشاف الصيغة العامة لتوزيع (t) عام 1908م، ونشر هذا الإنجـــاز تحت اسم مستعار هو (Student)، إلا أن اسم التوزيع بقي مقترناً باسم (Student) أي (t-Distribution)،

2-1- خصائص توزيع (t): يتميز توزيع (t) بالخصائص التالية:

- إن الوسط الحسابي (μ_x) و التباين (σ_x^2) والإنحراف المعياري (σ_x) لتوزيع (t) تكتب على النحو التالي:

$$\begin{cases} \mu_x = 0 \\ \sigma_x^2 = \dfrac{n}{n-2} \quad , n > 2 \\ \sigma_x = \sqrt{\dfrac{n}{n-2}} \end{cases}$$

والرسم الموالي يبين شكل منحنى التوزيع لدرجات حرية مختلفة:



يلاحظ من الرسم أن التوزيع متماثل حول النقطة $(t = 0)$، ويميل للإعتدال عند درجات الحرية الكبيرة $n \geq 30$.

وفيما يلي النظريات الهامة التي سنستخدمها في تطبيقاتنا.

نظرية: إذا كانت عينات عشوائية كل بحجم (n) وتباين (S^2) من مجتمع طبيعي التوزيع بوسط (μ) وتباين (σ^2) فإن:

$$T = \frac{X - \mu}{S / \sqrt{n-1}} \sim t_{(n-1)}$$

ونعبر عن قيمة المتغير العشوائي الذي يتبع توزيع (t) بالرمز $t_{(\alpha,\nu)}$ وهي القيمة التي يقع على يمينها مساحة (α) على منحنى توزيع (t) بدرجة حرية (ν)، وتحسب قيم هذا المتغير العشوائي والإحتمالات له باستخدام جدول توزيع (t).

ملاحظة: من تماثل التوزيع حول الصفر يكون: $t_{(\alpha,\nu)} = -t_{(1-\alpha,\nu)}$

مثال: أوجد القيم التالية: $t_{(0.995\ ,9)}$ ، $t_{(0.99\ ,11)}$ ، $t_{(0.05\ ,3)}$

الحل:

لإيجاد $t_{(0.05\ ,3)}$ نختار من العمود الأيسر (عمود درجة الحرية) القيمة 3، فنجد على يمينه صف من الأعداد ومن الصف العلوي (صف المساحات أي الإحتمالات) عن القيمة 0.05 فنجد أسفلها عمود من الأعـداد فتكـون القيمة الموجودة في تقاطع الصف والعمود هـي القيمـة المطلوبـة، وبالتـالي يكـون $t_{(0.05\ ,3)} = 2.353$.

– لإيجاد $t_{(0.99\ ,11)}$ نلاحظ أن جدول التوزيع لا يحتوي على القيمة $\alpha = 0.99$ لذلك نستغل تماثل منحنى االتوزيع حول الصفر فيكون: $t_{(0.99\ ,11)} = -t_{(0.99\ ,11)} = -2.718$

– بالمثل يكون $t_{(0.995\ ,9)} = -t_{(0.995\ ,9)} = -3.250$

2-2- استخدامات توزيع (t): إن لتوزيع (t) استخدامات كثيرة منها:

– اختبار متوسط مجتمع طبيعي مجهول التباين؛

– اختبار الفرق بين متوسطي مجتمعين طبيعيين مجهولي التباين؛

– اختبار معنوية معامل الإرتباط البسيط؛

– اختبار معنوية معاملات الإنحدار في الإنحدار الخطي المتعدد؛

– اختبار معنوية معامل الارتباط الجزئي؛

– تكوين فترات الثقة لمتوسط مجتمع طبيعي مجهول التباين.

3- توزيع(F): F-distribution:

تُنسب فكرة توزيع (F) إلى العالم الإحصائي (R.A.Fisher) خلال العقد الثاني من القرن العشرين، والـذي يُعد من أوائل الإحصائيين الذين وضعوا الأسس النظرية لهذا التوزيع خلال الفترة (1890-1962م)، وجاء من بعده العالم الإحصائي (Snedecor) الذي كان له الفضل الكبير في تطور هذا التوزيع.

إن الأساس النظري لاشتقاق توزيع (F) هو في الحقيقة توزيع مشتق من نسبة توزيعين مستقلين كل منهما عبارة عن توزيع كاي مربع (χ^2) مقسوماً على درجة حريته (df).

وعلى افتراض كان لدينا المتغير (χ_1^2) كمتغير عشوائي يخضع إلى توزيع كاي مربع بدرجات حرية (n − 1) حيث أن:

$$\chi_1^2 = \frac{(n-1)S_1^2}{\sigma_1^2} \sim \chi_{(n-1)}^2$$

وكان لدينا المتغير (χ_2^2) كمتغير عشوائي يخضع إلى توزيع كاي مربع بدرجات حرية (m − 1) حيث أن:

$$\chi_2^2 = \frac{(n-1)S_2^2}{\sigma_2^2} \sim \chi_{(m-1)}^2$$

3-1- خصائص توزيع (F): يتميز توزيع (F) بالخصائص التالية:

- إن قيمة (F) موجبة دائما أي أن: $(\mathrm{F}>0)$؛
- إن منحنى توزيع (F) ملتوٍ نحو اليمين (التواء موجب)؛
- إذا كانت $\mathrm{F}\sim \mathrm{F}_{(\nu_1,\nu_2)}$ فإن $\frac{1}{\mathrm{F}}\sim \mathrm{F}_{(\nu_2,\nu_1)}$

حيث أن: ν_1: درجة حرية البسط؛

ν_2: درجة حرية المقام.

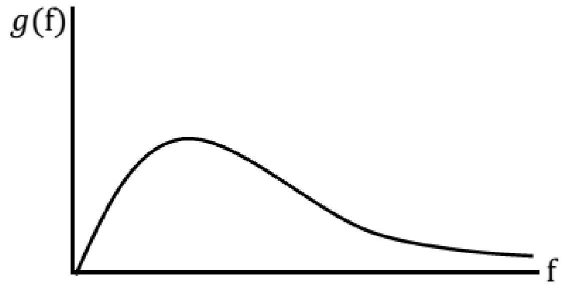
مثال: أوجد قيمة $\mathrm{F}_{(,7,5,0.95)}$

الحل:

لاحظ أن $1-\alpha=0.95$ ، و هي ليس من ضمن القيم المدرجة في جداول التوزيـــع (F)، لـــذلك بـــالتطبيق المباشر لهذه النظرية الأخيرة نجد أن:

$$\mathrm{F}_{(7,\,5,0.95,)}=\frac{1}{\mathrm{F}(0.95,\,5,\,7)}=\frac{1}{3.97}=0.253$$

وتأسيساً على ما سبق يعد توزيع (F) من أهم التوزيعات الإحتمالية المتصلة لما لها من أهميـــة في الكـــثير مـــن التطبيقات الإحصائية، ومن جانب آخر يعد هذا التوزيع أساساً نظرياً لكثير من الإختبارات الإحصائية المســـتندة على اختبار (F) الذي يتميز بتعدد استخداماته في الحياة العملية.



ونعبر عن قيمة المتغير العشوائي الذي يتبع توزيع (F) بالرمز $\mathrm{F}_{(\alpha,\nu_1,\nu_2)}$، و هي القيمة التي يقع علـــى يمينـــها مساحة (α) على منحنى توزيع (F) بدرجة حرية (ν_1) في البسط و(ν_2) في المقام.

مثال(1): أوجد القيم: $F_{(0.05, 1, 7)}$، $F_{(0.025, 3, 5)}$، $F_{(0.1, 12, 8)}$

الحل:

لحساب $F_{(0.05, 1, 7)}$ نختار من جداول التوزيع الجدول الخاص بقيمة $\alpha = 0.05$، ومن الصف في أعلى الجـــدول الخاص بدرجات حرية البسط نختار القيمة 1؛ حيث يوجد أسفلها عمود من الأعداد، ومن العمود أقصى اليسار والخاص بدرجة المقام نختار القيمة 7؛ حيث يوجد عل يمينه صف من الأعداد، وبتقاطع هذا الصف مع العمـــود أسفل درجة حرية البسط نجد أن: $F_{(0.05, 1, 7)} = 5.59$.

وبالمثل نجد أن: $F_{(0.025, 3, 5)} = 7.76$، $F_{(0.1, 12, 8)} = 2.50$

مثال(2): إذا كان $F_{(\alpha, 9, 8)} = 4.35$ أوجد قيمة α؟.

الحل:

بالبحث في جداول توزيع (F) و في كل حالات(α) عن القيم الموجودة في تقاطعات أعمدة درجة الحريـــة 9 الخاصة بالبسط مع القيم الموجوجة في صفوف درجة الحرية 8 الخاصة بالمقام نجد ما يلي:

α	0.1	0.05	0.025	0.01
$F_{(\alpha,9,8)}$	2.56	3.39	4.36	5.91

لاحظ أن القيمة 4.3572 هي أقرب قيمة من قيمة المتغير العشوائي المعطاة و هي 4.3452 للك نأخذ القيمة (α) تساوي 0.025.

في بعض الأحيان تكون قيمة (α) هي إحدى القيم {0.9، 0.95، 0.975، 0.99}، وهي ليس من ضمن القيم المدرجة في جداول توزيع (F).

3-2- استخدامات توزيع (F): إن لتوزيع (F) استخدامات كثيرة منها:

- اختبار تجانس تبايني عينتين مستقلتين؛
- اختبار معنوية معامل الإرتباط المتعدد؛
- اختبار معنوية نموذج الإنحدار الخطي المتعدد؛
- اختبار تجانس عدة تقديرات مستقلة لتباين مجتمع طبيعي؛
- الإستخدام الواسع في أسلوب تحليل التباين.