الفصل السابع

توزيعات المعاينة

Sampling Distributions

(٧-١) مقدمة Introduction

يهتم فرع الإحصاء الاستدلالى بالتعميم والتنبؤ، فعلى سبيل المثال يمكن القول أن متوسط دخل الفرد في بلد ما $ 86000 في السنة وذلك بناء على عينة عشوائية اختيرت من هذا البلد. وبتوضيح آخر يمكن أن نتوقع بناء على أراء مجموعة من الأشخاص في الشارع أن 80% من أصوات الناخبين في مدينة ما سوف تعطى لمرشح معين. كما يمكن التوقع أن عمر المصباح الكهربائى من إنتاج مصنع ما يتراوح بين 1150 ساعة و 1250 ساعة بدرجة ثقة معينة. فإننا نجد فى كل مثال من الأمثلة السابقة، تم حساب إحصاء من عينة عشوائية تم اختيارها من المجتمع موضع الدراسة، ومن تلك الإحصاءات أمكننا الوصول إلى جمل تخص قيم المعالم والتى قد تكون صحيحة أو غير صحيحة. التعميم من الإحصاء إلى المعلمة يكون بثقة فقط إذا استطعنا أن نفهم العلاقة بين المجتمع وعيناته.

في الفصل الخامس عرفنا الإحصاء على أنه متغير عشوائى يعتمد قيمته فقط على العينة، وبالتالي فإن نفس الحسابات لعينات مختلفة من المجتمع تودى إلى قيم مختلفة للإحصاء. هذه الاختلافات في قيم الإحصاء تعتمد على حجم المجتمع وحجم العينات والطريقة التى استخدمت في اختيار العينات العشوائية. إذا كان حجم المجتمع كبيراً أو لانهائي فإن التوزيع الاحتمالى للإحصاء في حالة السحب بإرجاع سوف يكون نفسه في حالة السحب بدون إرجاع. ومن ناحية أخري فإن السحب بإرجاع من مجتمع صغير محدود يعطى توزيعاً للإحصاء يختلف قليلاً عن السحب بدون إرجاع. أخيرا المعاينة مع الإرجاع من مجتمع محدود يكافئ المعاينة من مجتمع لانهائى وذلك لعدم وجود حدود لحجم العينة المختارة من المجتمع.

تعريف : التوزيع الاحتمالي لأي إحصاء يسمى التوزيع العينى **sampling distribution**

تعريف : الانحراف المعياري للتوزيع العينى لأى إحصاء يسمى الخطأ المعيارى **standard error** للإحصاء .

فعلى سبيل المثال التوزيع الاحتمالى للإحصاء $\overline{X}$ يسمى التوزيع العينى للمتوسط، كما أن الخطأ المعيارى للمتوسط هو الانحراف المعيارى للتوزيع العينى للإحصاء $\overline{X}$.

في هذا الفصل سوف ندرس بعض توزيعات المعاينة الأكثر استخداما في الإحصاء. التطبيقات على تلك التوزيعات العينية تخص مشاكل الاستدلال الإحصائى التى سوف نتناولها في الفصل الثامن والتاسع.

(٧-٢) توزيعات المعاينة الطبيعية Normal Sampling Distributions

إذا أخذنا عينات متكررة من الحجم n من توزيع متصل له متوسط μ وبتبـــاين σ^2. لكل عينة تم حساب القيمة y لإحصاء ما Y، والذي نفسه متغير عشوائي متصل. بفـــرض أن Y يتبع توزيعاً طبيعياً بمتوسط μ_Y وانحراف معياري σ_Y. النظرية التالية تنص على أن :

نظرية (٧-١) إذا كان y قيمة للإحصاء Y والذي يتبع توزيعاً طبيعياً بمتوسط μ_Y والانحراف معياري σ_Y، فإن :

$$z = \frac{y - \mu_Y}{\sigma_Y}$$

هي قيمة لمتغير عشوائي Z يتبع التوزيع الطبيعي القياسي حيث :

$$Z = \frac{Y - \mu_Y}{\sigma_Y}.$$

يمكن تطبيق نظرية (٧-١) للإحصاءات المحسوبة من عينات عشـــوائية اختـــيرت مـــن مجتمعات متقطعة، سواء محدودة أو غير محدودة، والتي التوزيعات العينية لإحصاءاتها تقريبـــاً تتبع توزيعاً طبيعياً. هذا ويمكننا استخدام جدول التوزيع الطبيعي في الملحق (٣) في حســـاب الاحتمالات التي يأخذها الإحصاء في فترات معينة.

(٧-٣) توزيعات المعاينة للمتوسط Sampling Distributions of the Mean

يعتبر توزيع المعاينة للمتوسط $\overline{X}$ أهم توزيع معاينة سوف نتناوله في هذا الفصـــل. إن شكل ونوع التوزيع الاحتمالي لمجتمع متوسط العينات (التوزيع العيني للمتوســـط) يعتمـــد على شكل المجتمع الأصلي الذى اختيرت منه العينات. النظرية الآتية تعطي التوزيع العيـــنى للمتوسط إذا كان المجتمع الأصلي التي اختيرت منه العينات يتبع توزيعاً طبيعيـــاً ونذكرهـــا بدون برهان.

نظرية (٧-٢) إذا أخذنا عينات متكررة من مجتمع معروف أنه يتبع توزيعـــاً طبيعيـــاً بمتوسط μ و انحراف معياري σ فإن التوزيع العينى للإحصاء $\overline{X}$ يتبع توزيعاً طبيعياً بمتوسط $\mu_{\overline{X}} = \mu$ وانحراف معياري $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$ حيث $\mu_{\overline{X}}$ و $\sigma_{\overline{X}}$ يرمزان للمتوسط والانحراف المعياري على التوالي للتوزيع العينى للإحصاء $\overline{X}$.

مثال (٧-١) إذا كانت أوزان الطلاب في جامعة ما تتبع توزيعاً طبيعياً متوسطة 70 كيلـــو جراما وانحرافه المعياري 10 كيلو جراما. اختيرت عينه عشوائية مكونة من 25 طالبا فما هو

احتمال أن يكون متوسط الأوزان أقل من 75 كيلو جرام.

الحل .

$$\mu = 70 \quad , \quad \sigma = 10 \quad , \quad n = 25$$

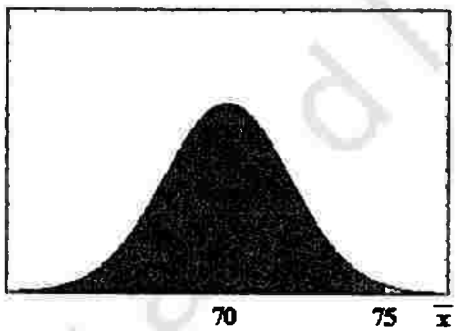
تبعا لنظرية (٧-١) فإن $\overline{X}$ يتبع توزيعاً طبيعياً متوسطة $\mu_{\overline{X}} = \mu = 70$ وانحرافه المعياري :

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{10}{\sqrt{25}} = 2.$$

والمطلوب هو حساب الاحتمال :

$$P(\overline{X} < 75).$$

والذي يساوى المساحة المظللة في شكل (٧-١).



شكل (٧-١)

عندما $\bar{x} = 75$ فإن قيمة z المقابلة لها هي :

$$z = \frac{\bar{x} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{75-70}{2} = 2.5$$

وعلى ذلك :

$$\begin{aligned} P(\overline{X} < 75) &= P(Z < 2.5) \\ &= 1 - P(Z > 2.5) \\ &= 0.5 + P(0 < Z < 2.5) \\ &= 0.5 + 0.4938 = 0.9938. \end{aligned}$$

في كثير من الحالات يكون التوزيع الاحتمالى للمجتمع الأصلي غير طبيعى ويتطلب الأمر معرفة التوزيع الاحتمالى للوسط الحسابي $\overline{X}$. للتسهيل سوف نستخدم مجتمعا منفصلا

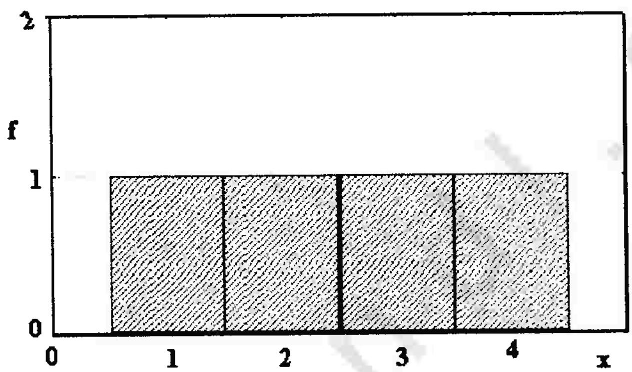
منتظما يتكون من القيم 1, 2, 3, 4 والذى متوسطة :

$$\mu = \frac{1+2+3+4}{4} = 2.5$$

وانحرافه المعياري :

$$\sigma = \sqrt{\frac{(1-2.5)^2 + (2-2.5)^2 + (3-2.5)^2 + (4-2.5)^2}{4}} = 1.118.$$

المدرج التكراري لهذا المجتمع موضح في شكل (٧-٢).



شكل (٧-٢)

بفرض أنه تم اختيار كل العينات من الحجم n = 2 من هذا المجتمع بإرجاع والذى يكافئ المعاينة من مجتمع لانهائي. يعطى جدول (٧-١) كل العينات الممكنة التى يمكن اختيارها (عدد العينات $N^n = 4^2 = 16$ عينة) من هذا المجتمع مع قيمها. لكل عينه تم حساب $\bar{x}$ والتوزيع التكراري لمجتمع متوسط العينات التى حجم كل منها n= 2 معطى في جدول (٧-٢) وتوزيعه التكراري في شكل (٧-٣).

جدول (٧-١)

رقم العينة	1	2	3	4	5	6	7	8
القيم	1,1	1,2	1,3	1,4	2,1	2,2	2,3	2,4
رقم العينة	9	10	11	12	13	14	15	16
القيم	3,1	3,2	3,3	3,4	4,1	4,2	4,3	4,4

جدول (٧-٢)

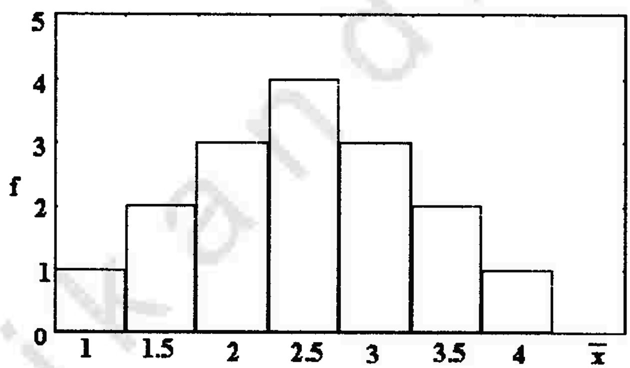
$\bar{x}$	1	1.5	2	2.5	3	3.5	4
f التكرار	1	2	3	4	3	2	1

يلاحظ أن توزيع المعاينة للإحصاء $\bar{X}$ في شكل (٧ -٣) تقريبا طبيعي. المتوسط و الانحراف المعياري للتوزيع العيني للإحصاء $\bar{X}$ تم حسابهما من جدول (٧ -٢) وهما على التوالي :

$$\mu_{\bar{X}} = \frac{\Sigma f \bar{x}}{\Sigma f} = \frac{40}{16} = 2.5 = \mu \ ,$$

$$\sigma_{\bar{X}} = \sqrt{\frac{\Sigma f (\bar{x} - 2.5)^2}{\Sigma f}} = \sqrt{\frac{10}{16}}$$

$$= 0.791 = \frac{1.118}{\sqrt{2}} = \frac{\sigma}{\sqrt{n}}.$$



شكل (٧-٣)

دائما المتوسط للإحصاء $\bar{X}$ يساوى متوسط المجتمع الذى اختيرت منه العينات العشوائية ولا يعتمد على حجم العينة. بينما الانحراف المعيارى للإحصاء $\bar{X}$ يعتمد على حجم العينة ويساوى الانحراف المعياري للمجتمع الأصلي σ مقسوما على $\sqrt{n}$. وعلى ذلك كلما زادت حجم العينة كلما قل الخطأ المعيارى للإحصاء $\bar{X}$ واقتربت $\bar{x}$ من μ وعلى ذلك يمكننا استخدام $\bar{x}$ كتقدير للمعلمة μ.

نظرية (٧ - ٣) إذا اختيرت كل العينات الممكنة من الحجم n بإرجاع من مجتمع محدود

من الحجم N وله متوسط μ وانحراف معياري σ فإن التوزيع العينى للإحصاء $\overline{X}$ تقريباً يتبع توزيعاً طبيعياً بمتوسط $\mu_{\overline{X}} = \mu$ وانحراف معياري $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$ وعلى ذلك :

$$z = \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

هي قيمة لمتغير عشوائى Z يتبع التوزيع الطبيعى القياسى . نظرية (٧-٣) صحيحة لأي مجتمع محدود عندما $n \geq 30$.

مثال (٧ - ٢) مجتمع مكون من المفردات الآتية :

$$2\ ,\ 2\ ,\ 2\ ,4\ ,5\ ,\ 7\ ,7\ ,\ 7,\ 8$$

أوجد احتمال أن عينة عشوائية من الحجم $n = 35$ اختيرت من هذا المجتمع بإرجاع تعطى متوسط عينة أكبر من 5.

الحل . المتوسط والانحراف المعياري للمجتمع المعطى هما :

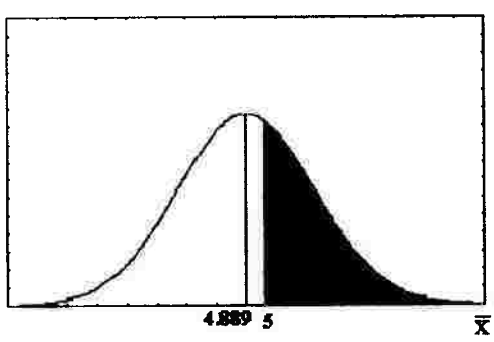
$$\mu = 4.889 \quad , \quad \sigma = 2.331 .$$

وحيث أن $n>30$ وتبعا لنظرية (٧-٣) فإن التوزيع العيني للإحصاء $\overline{X}$ تقريباً يتبع التوزيع الطبيعي بمتوسط $\mu_{\overline{X}} = \mu = 4.889$ وانحراف معياري $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{2.331}{\sqrt{35}} = 0.394$. احتمال أن $\overline{X}$ أكبر من 5 يساوى المساحة المظللة في شكل (٧-٤). قيمة z المقابلة لقيمة $\overline{x} = 5$ هي :

$$z = \frac{\overline{x} - \mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{5 - 4.889}{0.394} = 0.282.$$

وعلى ذلك :

$$\begin{aligned} P(\overline{X} > 5) &= P(Z > 0.282) \\ &= 0.5 - P(0 < Z < 0.282) . \\ &= 0.5 - 0.1103 = 0.3897. \end{aligned}$$



شكل (٧-٤)

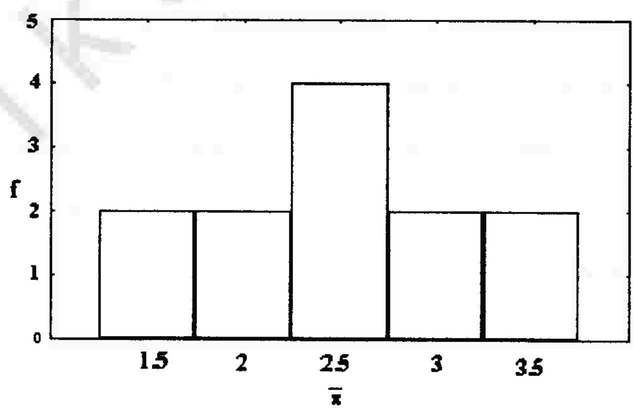
بفرض أننا سحبنا كل العينات الممكنة من الحجم $n = 2$ من مجتمعنا المنتظم والذى مشاهداته $1, 2, 3, 4$ ولكن بدون إرجاع. لكل عينة تم حساب متوسط العينة $\bar{x}$. يعطى جدول (٧-٣) كل العينات الممكنة التي يمكن اختيارها من هذا المجتمع بدون إرجاع مع قيم كل عينة (عدد العينات $12 = \frac{4!}{2!} = \frac{N!}{(N-n)!}$ عينة). التوزيع التكراري لمجتمع متوسط العينات من الحجم $n = 2$ معطى في جدول (٧-٤). المدرج التكراري لمجتمع متوسط العينات موضح في شكل (٧-٥).

جدول (٧-٣)

رقم العينة	1	2	3	4	5	6
القيم	1,2	1,3	1,4	2,1	2,3	2,4
رقم العينة	7	8	9	10	11	12
القيم	3,1	3,2	3,4	4,1	4,2	4,3

جدول (٧ - ٤)

$\bar{x}$	1.5	2.0	2.5	3.0	3.5
f	2	2	4	2	2



شكل (٧-٥)

يتضح من شكل (٧-٥) أن التوزيع العينى للإحصاء $\bar{X}$ في حالة السحب بدون إرجاع من مجتمع محدود بعيدا عن التوزيع الطبيعى حيث $n = 2$. من جدول (٧-٤) يمكن

حساب المتوسط والانحراف المعياري للإحصاء $\overline{X}$ على التوالي :

$$\mu_{\overline{X}} = \frac{\Sigma f \,\bar{x}}{\Sigma f} = \frac{30}{12} = 2.5 = \mu \;,$$

$$\sigma_{\overline{X}} = \sqrt{\frac{\Sigma f (\bar{x} - 2.5)^2}{\Sigma f}} = \sqrt{\frac{5}{12}} = 0.645$$

$$= \frac{1.118}{\sqrt{2}} \sqrt{\frac{4-2}{4-1}} = \frac{\sigma}{\sqrt{\mathbf{n}}} \sqrt{\frac{\mathbf{N}-\mathbf{n}}{\mathbf{N}-1}} = 0.645 .$$

عندما يكون $n \geq 30$ يمكن تطبيق النظرية التالية :

نظرية (٧-٤) إذا اختيرت كل العينات الممكنة بدون إرجاع من مجتمع محدود من الحجم N وله متوسط μ وانحراف معياري σ، فإن التوزيع العينى للإحصاء $\overline{X}$ تقريبا يتبع توزيعـــا طبيعيا بمتوسط وانحراف معياري معطى كالتالي :

$$\mu_{\overline{X}} = \mu,$$

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{\mathbf{n}}} \sqrt{\frac{\mathbf{N}-\mathbf{n}}{\mathbf{N}-1}}.$$

يسمى المقدار $\sqrt{\frac{\mathbf{N}-\mathbf{n}}{\mathbf{N}-1}}$ معامل التصحيح. إذا كان حجم العينة صغيرا جدا بالنسبة لحجم المجتمع فإن $(N-n) / (N-1)$ تكون قريبة من 1 ويمكن إسقاطها من المعادلة. وقـــد جـــرت العادة على إهمال هذا الحد عندما تكون $n < 0.05\,N$.

مثال (٧-٣) من مجتمع مكون من القيم

2, 2, 2, 4, 5, 6, 7, 7, 7, 8

المطلوب : (أ) حساب متوسط المجتمع الأصلى μ وانحرافه المعيارى σ.

(ب) حساب متوسط مجتمع متوسطات العينات $\mu_{\overline{X}}$ وانحرافـــه المعيـــاري $\sigma_{\overline{X}}$ عند $n=2$ (السحب بدون إرجاع).

الحل .

(أ) $\mu = 5, \; \sigma = 2.24 \,, \; N = 10, \; n = 2$

(ب) $\mu_{\overline{X}} = \mu = 5$

وحيث أن :

$$0.05\,N = (0.05)\,(10) = 0.5$$

أى أن $n = 2$ سوف تكون أكبر من 0.5 $(n > .05\,N)$ وعلى ذلك لا يمكننا إهمال معامل التصحيح في صيغة $\sigma_{\overline{X}}$. وعلى ذلك :

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{2.24}{\sqrt{2}}\sqrt{\frac{10-2}{10-1}} = 1.4933.$$

مثال (٧-٤) افترض مجتمعا ما يتكون من 1000 عنصر له متوسط حسابي $\mu=15$ وانحراف معياري $\sigma = 6$ أوجد :

(أ) المتوسط الحسابي والانحراف المعيارى لتوزيع المعاينة للإحصاء $\overline{X}$ عندما يكون $n = 36$ (السحب بدون إرجاع).

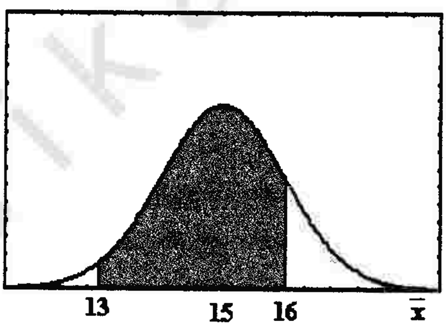
(ب) احتمال أن يقع متوسط العينة العشوائية من الحجم $n = 36$ بين 13, 16.

الحل .

(أ) بما أن $n = 36 < 0.05\,N = 50$ فإن التوزيع العينى للإحصاء $\overline{X}$ يكون له متوسط وانحراف معيارى كالتالي :

$$\mu_{\overline{X}} = \mu = 15,$$
$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{6}{\sqrt{36}} = 1.$$

(ب) الاحتمال المطلوب موضح بالمساحة المظللة في شكل (٧-٦) .



شكل (٧-٦)

قيمة z_1 المقابلة لقيمة $\overline{x}_1=13$ هي :

$$z_1 = \frac{\overline{x}_1 - \mu}{\sigma_{\overline{x}}} = \frac{13-15}{1} = -2,$$

قيمة z_2 المقابلة للقيمة $\overline{x}_2 = 16$ هي :

$$z_2 = \frac{\bar{x}_2 - \mu}{\sigma_{\bar{x}}} = \frac{16-15}{1} = 1.$$

وعلى ذلك :

$$P(13 < \bar{X} < 16) = P(-2 < Z < 1)$$
$$= P0 < Z < 2) + P(0 < Z < 1)$$
$$= 0.4772 + 0.3413 = 0.8185.$$

للمجتمعات الكبيرة أو اللانهائية سواء كانت متصلة أم متقطعة تنص النظرية التالية على .

نظرية (٧-٥) إذا اختيرت كل العينات الممكنة من الحجم n من مجتمع كبير أو لانهائي بمتوسط μ وتباين σ^2 فإن التوزيع العينى للإحصاء $\bar{X}$ تقريبا يتبع توزيعا طبيعيا بمتوسط :

$$\mu_{\bar{X}} = \mu$$

وانحراف معياري :

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}.$$

وعلى ذلك فإن :

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

هي قيمة لمتغير عشوائى Z يتبع التوزيع الطبيعى القياسي.

التقريب الطبيعي في نظرية (٧-٥) سوف يكون جيدا إذا كانت $n \geq 30$ بصرف النظر عن شكل المجتمع الأصلى الذى اختيرت منه العينات. إذا كانت $n<30$ التقريب سوف يكون جيد فقط إذا كان المجتمع لا يختلف كثيرا عن التوزيع الطبيعي.

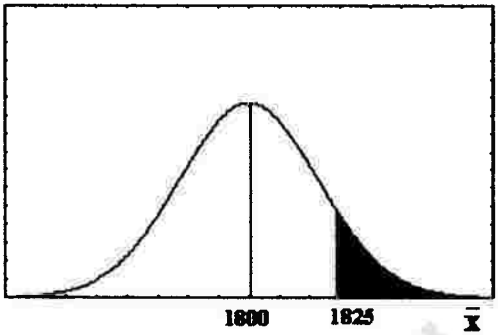
مثال (٧-٥) إذا كانت أعمار المصابيح المنتجة بواسطة أحد المصانع لها متوسط عمر $\mu = 1800$ ساعة بانحراف معيارى $\sigma = 200$ ساعة. أوجد احتمال أن عينة عشوائية من 100 مصباح سوف يكون لها متوسط عمر أكبر من 1825 ساعة.

الحل .

المجتمع كبير والعينة هنا كبيرة n=100، على ذلك التوزيع العينى للإحصاء $\bar{X}$ تقريبا يتبع توزيعا طبيعيا بمتوسط 1800 ساعة وانحراف معياري $20 = \frac{200}{\sqrt{100}} = \frac{\sigma}{\sqrt{n}}$. وعلى ذلك عندما يكون $\bar{x} = 1825$ فإن :

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{1825 - 1800}{20} = 1.25.$$

الاحتمال المطلوب موضح بالمساحة المظللة في شكل (٧-٧).



شكل (٧-٧)

وعلى ذلك :

$$P(\bar{X} > 1825) = P(Z > 1.25)$$
$$= 0.5 - 0.3944 = 0.1056.$$

مثال (٧-٦) إذا كان متوسط الدب الأسترالي $\mu = 20$ بوصة بانحراف معياري $\sigma = 4$ بوصة. فإذا خطط لاختيار عينة عشوائية من الحجم $n = 64$. ما هو احتمال أن يكون متوسط العينة أكبر من 21 بوصة.

الحل . التوزيع العيني للإحصاء $\bar{X}$ تقريبا يتبع توزيعا طبيعيا بمتوسط $\mu = 20$ وانحراف معياري :

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{4}{\sqrt{64}} = 0.5.$$

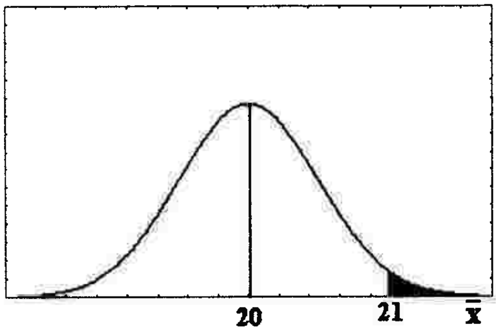
احتمال أن يكون متوسط عينه عشوائية من الحجم $n = 64$ أكبر من 21 بوصة يساوى المساحة المظللة في شكل (٧-٨).

عندما يكون $\bar{x} = 21$ فإن :

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{X}}} = \frac{21 - 20}{0.5} = 2.$$

وعلى ذلك :

$$P(\overline{X} > 21) = P(Z > 2) = 0.5 - P(0 < Z < 2)$$
$$= 0.5 - 0.4772 = 0.0228.$$



شكل (٧–٨)

مثال (٧–٧) إذا كان متوسط سمك الرخام المنتج في مصنع للعب الأطفال هو 0.85سم بانحراف معياري 0.01 المطلوب :

(أ) احتمال اختيار عينة عشوائية من 100 قطعة رخام لها سمك أكبر من 0.851.

(ب) ما هما القيمتان التى تتوقع أن يكون %95 من متوسطات العينات بينهما .

الحل .

(أ) التوزيع العينى للإحصاء $\overline{X}$ تقريبا يتبع توزيعا طبيعيا بمتوسط $\mu_{\overline{X}} = \mu = 0.85.$ وانحراف معياري :

$$\frac{\sigma}{\sqrt{n}} = \frac{0.01}{\sqrt{100}} = 0.001.$$

عندما يكون $\overline{x} = 0.851$ فإن :

$$z = \frac{\overline{x} - \mu}{\sigma_{\overline{X}}} = \frac{0.851 - 0.85}{0.001} = 1.0.$$

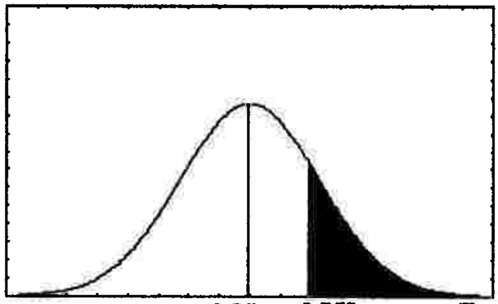
الاحتمال المطلوب يساوي المساحة المظللة في شكل (٧–٩) وعلى ذلك :

$$P(\overline{X} > 0.851) = P(Z > 1.0)$$
$$= 0.5 - 0.3413 = 0.1587.$$

(ب) باستخدام القاعدة التجريبية فإننا نتوقع أن % 95 متوسطات العينات تقع في الفترة :

$$\mu_{\overline{X}} \pm 2\sigma_{\overline{X}} = 0.85 \pm 2(0.001)$$

أي في الفترة (0.848, 0.852).



شكل (٧-٩)

(٧-٤) التوزيعات العينية للفرق بين متوسطى مجتمعين

Sampling Distributions of the Different Between Two Populations Means

بفرض أن لدينا مجتمعين الأول متوسطة μ_1 وتباينه σ_1^2 والثاني متوسطة μ_2 وتباينه σ_2^2. بفرض أن قيم المتغير $\overline{X}_1$ تمثل متوسطات لعينات عشوائية من الحجم n_1 اختيرت من المجتمع الأول، وقيم المتغير $\overline{X}_2$ تمثل متوسطات لعينات عشوائية من المجتمع الثاني ومستقلة عن المجتمع الأول. التوزيع للفروق $\bar{x}_1 - \bar{x}_2$ بين الفئتين من متوسطات العينتين المستقلتين يسمى التوزيع العينى للإحصاء $\overline{X}_1 - \overline{X}_2$. للتسهيل نفرض أن المجتمع الأول من الحجم $N_1 = 3$ ويتكون من القيم 4, 5, 6 والتى متوسطها :

$$\mu_1 = \frac{4+5+6}{3} = 5$$

وتباينها :

$$\sigma_1^2 = \frac{(4-5)^2 + (5-5)^2 + (6-5)^2}{3} = \frac{2}{3}.$$

المجتمع الثاني يتكون من القيمتين 1, 4 ولهما المتوسط :

$$\mu_2 = \frac{1+4}{2} = 2.5$$

والتباين :

$$\sigma_2^2 = \frac{(1-2.5)^2 + (4-2.5)^2}{2} = \frac{9}{4}.$$

من المجتمع الأول تم اختيار كل العينات الممكنة من الحجم $n_1 = 2$ مع الإرجاع وحساب المتوسط $\bar{x}$ لكل عينة. بنفس الشكل للمجتمع الثاني تم اختيار كل العينات الممكنة من الحجم $n_2 = 3$ وحساب $\bar{x}_2$ لكل عينة. الفئتان من كل العينات ومتوسطاها معطاة في جدول (٧-٥).

جدول (٧-٥)

المجتمع الأول	رقم العينة	1	2	3	4	5	6	7	8	9
	القيم	4,4	4,5	4,6	5,4	5,5	5,6	6,4	6,5	6,6
	$\bar{x}_1$	4.0	4.5	5.0	4.5	5.0	5.5	5.0	5.5	6.0
المجتمع الثاني	رقم العينة	1	2	3	4	5	6	7	8	
	القيم	1,1,1	1,1,4	1,4,1	4,1,1	4,4,1	1,4,4	4,1,4	4,4,4	
	$\bar{x}_2$	1	2	2	2	3	3	3	4	

الفروق الممكنة والتي عددها 72 من $\bar{x}_1 - \bar{x}_2$ معطاة في جدول (٧-٦).

جدول (٧-٦)

$\bar{x}_2$	$\bar{x}_1$								
	4.0	4.5	5.0	4.5	5.0	5.5	5.0	5.5	6.0
1	3.0	3.5	4.0	3.5	4.0	4.5	4.0	4.5	5.0
2	2.0	2.5	3.0	2.5	3.0	3.5	3.0	3.5	4.0
2	2.0	2.5	3.0	2.5	3.0	3.5	3.0	3.5	4.0
2	2.0	2.5	3.0	2.5	3.0	3.5	3.0	3.5	4.0
3	1.0	1.5	2.0	1.5	2.0	2.5	2.0	2.5	3.0
3	1.0	1.5	2.0	1.5	2.0	2.5	2.0	2.5	3.0
3	1.0	1.5	2.0	1.5	2.0	2.5	2.0	2.5	3.0
4	0.0	0.5	1.0	0.5	1.0	1.5	1.0	1.5	2.0

التوزيع التكراري للإحصاء $\bar{X}_1 - \bar{X}_2$ في جدول (٧-٧) ومدرجة التكراري في شكل (٧-١٠). من الواضح أن المتغير العشوائي $\bar{X}_1 - \bar{X}_2$ تقريبا يتبع توزيعا طبيعيا وهذا

التقريب يتحسن عندما يزيد n_1, n_2. بتطبيق نظرية (٤-٦) ثم نظريــة (٧-٣) فــإن متوسط الفروق لمجتمع العينات المستقلة هو :

$$\mu_{\bar{X}_1-\bar{X}_2} = E(\bar{X}_1 - \bar{X}_2) = E(\bar{X}_1) - E(\bar{X}_2) = \mu_1 - \mu_2.$$

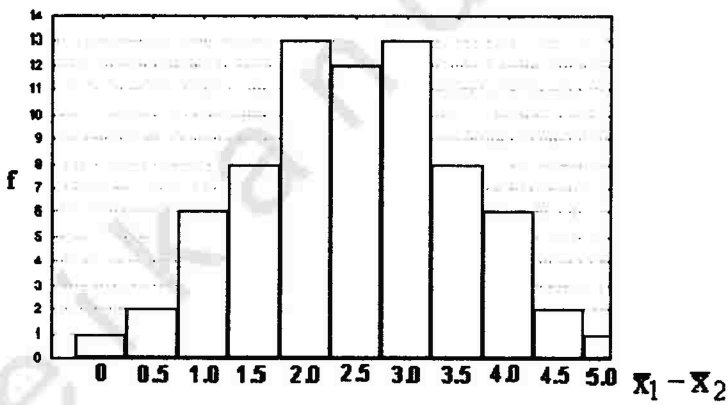
هذه النتيجة يمكن التحقق منها من البيانات في جدول (٧-٧) حيث :

$$\mu_{\bar{X}_1-\bar{X}_2} = \frac{\Sigma f(\bar{x}_1 - \bar{x}_2)}{\Sigma\ f} = \frac{180}{72} = 2.5 = 5 - 2.5$$

$$= \mu_1 - \mu_2.$$

جدول (٧-٧)

$\bar{x}_1 - \bar{x}_2$	0.0	.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
f	1	2	6	8	13	12	13	8	6	2	1



شكل (٧-١٠)

أيضا بتطبيق نظرية (٤-٨) ثم نظرية (٧-٣) فإن التباين لفروق المتوسطات المستقلة هو :

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

$$= \left[\left(\frac{2}{3}\right)/2\right] + \left[\left(\frac{9}{4}\right)/3\right] = 1.08333.$$

هذه النتيجة يمكن التحقق منها بسهولة وذلك بحساب التباين (1.08333) من البيانـــات في جدول (٧-٧). النتائج التي تم الحصول عليها للتوزيع العيني للإحصاء $\overline{X}_1 - \overline{X}_2$ عند المعاينة بإرجاع من مجتمع محدود تكون صحيحة للمجتمعات اللانهائية ســـواء المتقطعـــة أو المتصلة وأيضا للمجتمعات المحدودة عند المعاينة بدون إرجاع بشرط أن أحجام المجتمعات N_2 , N_1 تكون كبيرة نسبيا عن أحجام العينات n_2 ، n_1 على التوالي. أما إذا كان حجم المجتمع صغيرا والسحب بدون إرجاع فلابد من حساب $\sigma^2_{\overline{X}_2}, \sigma^2_{\overline{X}_1}$ من صيغـــة $\sigma_{\overline{X}}$ في نظرية (٧-٤).

تقتصر الدراسة من الآن وفي الفصول التالية على التوزيـــع العيـــنى للفـــروق بـــين المتوسطات المستقلة فقط إذا كان حجم المجتمع الذى تختار منه العينات كبيراً أو لانهائي.

نظرية (٧-٦) إذا اختيرت عينات مستقلة من الحجم N_2 ، N_1 من مجتمعين كـــبيرين (أو لانهائيتين) ، متقطعة أو متصلة، بمتوسطي μ_1, μ_2 وتبايني σ_1^2, σ_2^2 على التوالي، فإن التوزيع العينى لفروق المتوسطات، $\overline{X}_1 - \overline{X}_2$ ، تقريباً يتبـــع توزيعـــاً طبيعيـــاً بمتوســـط وانحراف معياري معطى كالتالي :

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

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

وعلى ذلك :

$$z = \frac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

هي قيمة لمتغير عشوائى Z يتبع التوزيع الطبيعى القياسى.

إذا كان كل من n_2 ، n_1 أكبر من أو يساوى 30 ، فإن التقريـــب الطبيعـــى لتوزيـــع $\overline{X}_1 - \overline{X}_2$ يكون جيداً جداً.

ملحوظة : إذا كانت العينتين تم اختيارهما من مجتمعين طبيعيين فإن :

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

هي قيمة لمتغير عشوائي Z يتبع التوزيع الطبيعي القياسي بصرف النظر عن حجم كلا من n_2 ، n_1 .

مثال (٧–٨) ينتج المصنع A بطاريات سيارة لها متوسط عمر 3.5 سنة بانحراف معياري 0.45 سنة. نفس البطاريات تنتج من المصنع B بمتوسط عمر 3.3 سنة وانحراف معياري 0.3 سنة. ما هو احتمال أن عينة عشوائية من 30 بطارية من المصنع A يكون لها متوسط عمر على الأقل يزيد 0.4 سنة عن متوسط عمر 36 بطارية من المصنع B ؟

الحل . سوف يكون لدينا البيانات التالية :

المجتمع الثاني	المجتمع الأول
$\mu_2 = 3.3$	$\mu_1 = 3.5$
$\sigma_2 = 0.3$	$\sigma_1 = 0.45$
$n_2 = 36$	$n_1 = 30$

العينات هنا كبيرة بدرجة كافية بحيث أن كل من $\overline{X}_2$ و $\overline{X}_1$ تقريبا يتبع توزيعا طبيعيا وعلى ذلك فإن $\overline{X}_1 - \overline{X}_2$ تقريبا تتبع توزيعا طبيعيا بمتوسط وانحراف معياري على التوالي :

$$\mu_{\overline{X}_1 - \overline{X}_2} = \mu_1 - \mu_2 = 3.5 - 3.3 = 0.2,$$

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

$$= \sqrt{\frac{0.45^2}{30} + \frac{0.3^2}{36}} = 0.096177,$$

الاحتمال المطلوب موضح بالمساحة المظللة في شكل (٧–١١).

عندما : $\bar{x}_1 - \bar{x}_2 = 0.4$

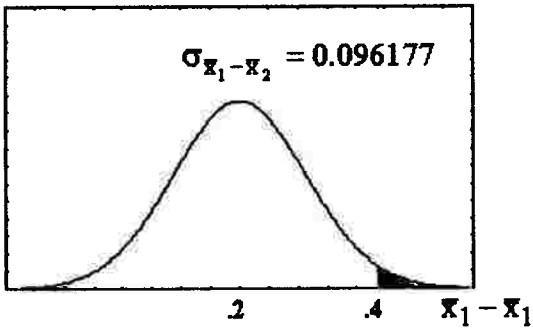
فإن :

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\overline{X}_1 - \overline{X}_2}} = \frac{0.4 - 0.2}{0.096177} = 2.08\,,$$

وعلى ذلك :

$$P(\overline{X}_1 - \overline{X}_2 > 0.4) = P(Z > 2.08)$$
$$= 0.5 - P(0 < Z < 2.08)$$

$$= 0.5 - 0.4812 = 0.0188.$$



شكل (٧-١١)

(٧-٥) التوزيعات العينية للنسب

Sampling Distributions of Proportions

بفرض أن لدينا مجتمعا ما وأن بعض مفردات هذا المجتمع تتوفر فيها صفة معينة وأن نسبة هذه المفردات هي p. فعلى سبيل المثال p قد تكون نسبة الأفراد المصابين بتسوس الأسنان في مدينة ما أو نسبة الطلبة المدخنين في كلية ما أو نسبة الوحدات المعيبة في مصنع ما. .. الخ . إذا أخذنا عينة عشوائية من الحجم n من هذا المجتمع ووجدنا من بينها x مفرده تتوفر فيها الصفة، وتم حساب $\hat{p} = \frac{x}{n}$ والتي تمثل نسبة المفردات في العينة والتي تتوفر فيها الصفة المعنية. إذا أخذنا عينات متكررة من الحجم n من هذا المجتمع فإن $\hat{p}$ تتغير من عينة إلى أخرى وتمثل قيمة للإحصاء $\hat{P}$. الآن سوف نتعرف على كيفية اشتقاق التوزيع العيني للنسبة وخصائصه سواء في حالة السحب بإرجاع أو بدون إرجاع وذلك من الأمثلة التالية:

مثال (٧-٩) مجتمع يتكون من القيم 1, 2, 3, 4 فإذا تم سحب كل العينات الممكنة من الحجم n = 2 من هذا المجتمع (بإرجاع). المطلوب إيجاد التوزيع العيني للإحصاء $\hat{P}$ والذي يمثل نسبة ظهور الرقم 4 في العينة. وإثبات أن المتوسط والتباين للتوزيع العيني للإحصاء $\hat{P}$ هما على التوالي :

$$\mu_{\hat{P}} = p,$$

$$\sigma^2_{\hat{P}} = \frac{pq}{n}.$$

الحل . الجدول (٧–٨) يحتوى على كل العينات الممكنة ونسبة ظهور الرقم 4 فيها.

جدول (٧–٨)

رقم العينة	القيم	عدد مرات ظهور 4	نسبة ظهور 4	رقم العينة	القيم	عدد مرات ظهور 4	نسبة ظهور 4
1	1,1	0	0	9	3,1	0	0
2	1,2	0	0	10	3,2	0	0
3	1,3	0	0	11	3,3	0	0
4	1,4	1	0.5	12	3,4	1	.5
5	2,1	0	0	13	4,1	1	.5
6	2,2	0	0	14	4,2	1	.5
7	2,3	0	0	15	4,3	1	.5
8	2,4	1	0.5	16	4,4	2	1

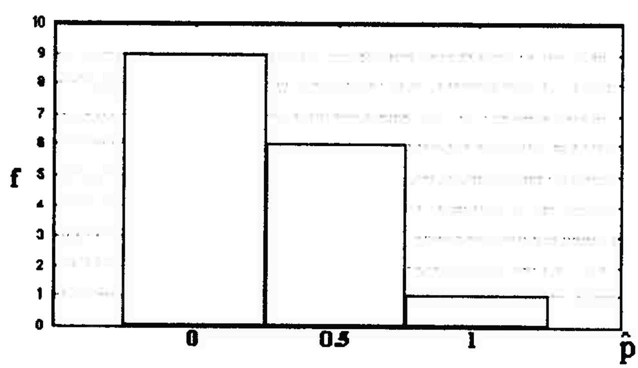
التوزيع التكراري لنسبة ظهور الرقم 4 للعينات من الحجم $n = 2$ التي تم اختيارها من المجتمع الذى حجمه $N = 4$ (بإرجاع) معطاة في جدول (٧–٩) ومدرجها التكراري في شكل (٧–١٢).

جدول (٧ –٩)

$\hat{p}$	0	0.5	1
f	9	6	1

التوزيع التكراري في شكل (٧ –١٢) ملتو ناحية اليمين وذلك لأن $p < 0.5$. إذا كانت $p > 0.5$ فإن التوزيع سوف يكون ملتويا ناحية اليسار .

شكل (٧–١٢)



المتوسط والتباين للتوزيع العيني للإحصاء $\hat{P}$ تم حسابهما من جدول (٧-٩) وهما :

$$\mu_{\hat{P}} = \frac{\Sigma f\, \hat{p}}{\Sigma f} = \frac{4}{16} = 0.25 = p,$$

$$\sigma^2_{\hat{P}} = \frac{\Sigma f(\hat{p} - 0.25)^2}{\Sigma f} = 0.094$$

$$= \frac{(0.25)(0.75)}{2} = \frac{\mathbf{pq}}{\mathbf{n}}.$$

مثال (٧-١٠) أوجد التوزيع العينى للإحصاء $\hat{P}$ للبيانات في مثال (٧-٩) إذا كان السحب بدون إرجاع وأثبت أن متوسط وتباين التوزيع العيني للإحصاء $\hat{P}$ هما على التوالي:

$$\mu_{\hat{P}} = p,$$

$$\sigma^2_{\hat{P}} = \frac{pq}{n} \cdot \left(\frac{N-n}{N-1}\right).$$

الجدول (٧-١٠) يحتوى على كل العينات ونسبة ظهور الرقم 4 فيها .

جدول (٧-١٠)

رقم العينة	القيم	عدد مرات ظهور 4	نسبة ظهور 4
1	1,2	0	0
2	1,3	0	0
3	1,4	1	.5
4	2,1	0	0
5	2,3	0	0
6	2,4	1	.5
7	3,1	0	0
8	3,2	0	0
9	3,4	1	.5
10	4,1	1	.5
11	4,2	1	.5
12	4,3	1	.5

التوزيع التكراري لنسبة ظهور الرقم 4 للعينات من الحجم $n = 2$ والتي تم اختيارها من المجتمع الذى حجمه $N = 4$ (بدون إرجاع) معطاة في جدول (٧-١١) ومدرجها التكراري في شكل (٧-١٣).

جدول (٧-١١)

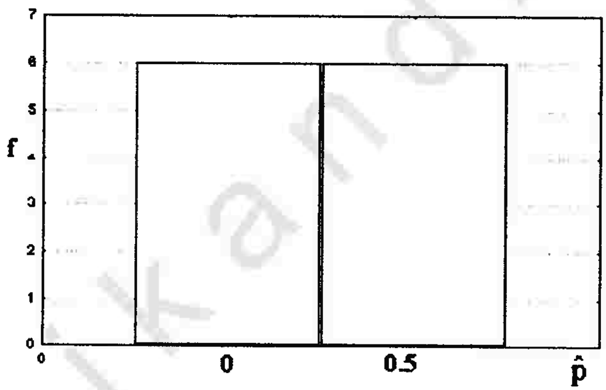
$\hat{p}$	0	0.5
f	6	6

المتوسط والتباين للتوزيع العيني للإحصاء $\hat{P}$ تم حسابهما من جدول (٧-١١) وهما :

$$\mu_{\hat{p}} = \frac{\Sigma f\, \hat{p}}{\Sigma f} = \frac{3}{12} = .25 = p,$$

$$\sigma^2_{\hat{p}} = \frac{\Sigma f(\hat{p} - 0.25)^2}{\Sigma f} = 0.0625.$$

$$= \frac{(0.25)(0.75)}{2}\left(\frac{4-2}{4-1}\right) = \frac{\mathbf{pq}}{\mathbf{n}}\left(\frac{\mathbf{N-n}}{\mathbf{N-1}}\right).$$



شكل (٧ -١٣)

يتضح من المثال السابق أن معامل التصحيح $\left(\frac{N-n}{N-1}\right)$ في صيغة $\sigma^2_{\hat{p}}$ يستخدم إذا كان المجتمع محدود والسحب بدون إرجاع. إذا كان حجم العينة أصغر من 0.05N يمكن اعتبار $\left(\frac{N-n}{N-1}\right) = 1$.

مثال (٧-١١) إذا كانت نسبة المدخنين في مجتمع من الحجم N= 350 هو 0.4.

سحبت عينة عشوائية من الحجم $n = 60$ من هذا المجتمع (بدون إرجاع) أوجد المتوسط والتباين للتوزيع العيني للإحصاء $\hat{P}$.

الحل .

$$\mu_{\hat{P}} = .4,$$

التباين يحسب من الصيغة :

$$\sigma^2_{\hat{P}} = \frac{pq}{n}\left(\frac{N-n}{N-1}\right).$$

وذلك لأن $n=60 > 0.05N=17.1$ ، وعلى ذلك :

$$\sigma^2_{\hat{P}} = \frac{pq}{n}\left(\frac{N-n}{N-1}\right) = \frac{(0.4)(0.6)}{60}\left(\frac{350-60}{350-1}\right) = 0.0033.$$

للمجتمعات الكبيرة أو اللانهائية فإن التوزيع العيني للإحصاء $\hat{P}$ تحدده النظرية التالية :

نظرية (٧-٧) إذا كانت p هى نسبة صفة معينة في مجتمع ما واختيرت من هذا المجتمع عينات كبيرة، حجم كل منها n وكان الإحصاء $\hat{P}$ يمثل نسبة وجود هذه الصفة في العينات فإن $\hat{P}$ تقريباً تتبع توزيعاً طبيعياً متوسطة وتباينه على التوالي :

$$\mu_{\hat{P}} = p,$$

$$\sigma^2_{\hat{P}} = \frac{pq}{n}.$$

وعلى ذلك :

$$z = \frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$$

هي قيمة لمتغير عشوائى Z تقريباً يتبع التوزيع الطبيعي القياسى.

وضع العالم Cochran (1963) قواعد لحجم العينة اللازم لتطبيق نظرية (٧-٧) معطاة في جدول (٧-١٢).

جدول (٧-١٢)

إذا كانت p تساوى	يستخدم التقريب الطبيعي إذا كان n على الأقل يساوى
.5	30
.4 - .6	50
.3 - .7	80
.2 - .8	200
.1 - .9	600
.05 - .95	1400

مثال (٧-١٢) بفرض أن مجتمعا ما أفراده عدة آلاف يمثل مصنعا لإنتاج كروت المعايدة. فإذا كان 0.2 من الكروت المنتجة تالفة (ا) أوجد المتوسط والانحراف للتوزيع العيني للإحصاء $\hat{P}$ وذلك عندما يكون $n = 300$ (ب) ما هو الاحتمال أن عينه عشوائية من الحجم $n = 300$ تعطى نسبة صفة أكبر من 0.19 (ج) أوجد احتمال أن يكون نسبة الكروت المنتجة التالفة تزيد عن 0.052 .

الحل .

(ا) بما أن المجتمع كبيراً و $n = 300, \ p = 0.2$ تحققان القواعد في جدول (٧-١٢)

وعلى ذلك فإن التوزيع العيني للإحصاء $\hat{P}$ تقريباً يتبع توزيعاً طبيعياً بمتوسط:

$$\mu_{\hat{p}} = p = 0.2$$

وانحراف معياري :

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.2)(0.8)}{300}} = 0.023.$$

(ب) عندما يكون $\hat{p} = 0.19$ فإن

$$z = \frac{\hat{p} - p}{\sigma_{\hat{p}}} = \frac{0.19 - 0.2}{0.023} = -0.43.$$

وعلى ذلك فإن :

$$\begin{aligned} P(\hat{P} > 0.19) &= P(Z > -0.43) \\ &= 0.5 + p\,(0 < Z < 0.43) \\ &= 0.5 + 0.1664 = 0.6664. \end{aligned}$$

(ج) عندما يكون $\hat{p} = 0.052$ فإن :

$$z=\frac{0.052-0.2}{0.023}=-6.43\,,$$

$$\mathbf{P\,(\hat{P}\;>\;0.052\,)=P\,(Z>\;-6.43)\approx 1}$$

مثال (٧-١٣) للمثال (٧-١١) أوجد احتمال أن تكون نسبة المدخنين في العينة أكبر من 0.35 .

الحل . عندما يكون $\hat{p}=0.35$ فإن :

$$z=\frac{0.35-0.4}{0.058}=-0.86$$

وعلى ذلك فإن :

$$\begin{aligned}\mathbf{P\,(\hat{P}\;>\;0.35\,)} &= \mathbf{P\,(Z>-\,0.86)}\\ &= \mathbf{0.5+P\,(\,0<Z<\;0.86\,)}\\ &= \mathbf{0.5\;+0.3051}\\ &= \mathbf{0.8051.}\end{aligned}$$

يمكن تطبيق نظرية (٧-٧) في حالة السحب بدون إرجاع من مجتمع محدود إذا كان حجم العينة أكبر من 0.05 N حيث متوسط التوزيع العيني للإحصاء $\hat{P}$ في هذه الحالة مازال $\mu_{\hat{P}}=p$ ولكن التباين يساوى :

$$\sigma^2_{\hat{P}}=\frac{\mathbf{pq}}{\mathbf{n}}\left(\frac{\mathbf{N-n}}{\mathbf{N-1}}\right).$$

وعلى ذلك فإن :

$$z=\frac{\hat{p}-p}{\sigma_{\hat{P}}}$$

هي قيمة لمتغير عشوائي Z يتبع التوزيع الطبيعي القياسي.

مثال (٧-١٤) إذا كانت نسبة المصابين بتسوس الأسنان في مجتمع من الحجم N= 200 هو 0.3. سحبت عينة عشوائية من الحجم n = 80 بدون إرجاع أوجد احتمال أن تكون نسبة المصابين بتسوس الأسنان في العينة أكبر من 0.35.

الحل . حجم العينة (n=80) أكبر من 0.05N=10 ، التوزيع العيني للإحصاء $\hat{P}$ سوف يكون تقريباً يتبع التوزيع الطبيعى لأن n = 80 تحقق القواعد في جدول (٧-١٢).

المتوسط و الانحراف المعياري للإحصاء $\hat{P}$ سوف يكون :

$$\mu_{\hat{P}}=\mathbf{p}=0.3\;,$$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}\sqrt{\frac{N-n}{N-1}}$$
$$= \sqrt{\frac{(.3)(.7)}{80}}\sqrt{\frac{200-80}{199}} = 0.03979.$$

عندما يكون $\hat{p} = 0.35$ فإن :

$$z = \frac{\hat{p}-p}{\sigma_{\hat{p}}} = \frac{0.35-0.3}{0.03979} \approx 1.26.$$

$$P(\hat{P} > .35) = P(Z > 1.26) = 0.5 - P(0 < Z\ 1.26)$$
$$= 0.5 - 0.3962 = 0.1038.$$

إذا كان لدينا مجتمعان مستقلان (مجتمعات كبيرة أو لانهائية) وإذا كانت p_1 هي نسبة توفر صفة ما في المجتمع الأول وكانت p_2 هى نسبة توفر الصفة نفسها في المجتمع الثانى . إذا اخترنا عينة عشوائية كبيرة حجمها n_1 من المجتمع الأول وحسبنا منها نسبة توفر الصفة محل الدراسة ولتكن $\hat{p}_1$. وإذا اخترنا عينة عشوائية كبيرة حجمها n_2 من المجتمع الثانى وحسبنا منها نسبة توفر الصفة المطلوبة ولتكن $\hat{p}_2$. بتكرار المعاينة من الحجم n_1 و n_2 فإن التوزيع العينى للإحصاء $\hat{P}_1 - \hat{P}_2$ تحدده النظرية التالية :

نظرية (٧-٨) التوزيع العينى للإحصاء $\hat{P}_1 - \hat{P}_2$ تقريباً يتبع توزيعاً طبيعياً بمتوسط :

$$\mu_{\hat{P}_1-\hat{P}_2} = p_1 - p_2,$$

وتباين :

$$\sigma^2_{\hat{P}} = \frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}.$$

وعلى ذلك تكون :

$$z = \frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}}$$

هي قيمة لمتغير عشوائى Z الذى يتبع التوزيع الطبيعى القياسى .

تعطى النظرية (٧-٨) نتائجاً جيدة ، إذا كانت n_1 , n_2 محددتان طبقا لقواعد Cochran المعطاة في جدول (٧-١٢) .

(٧-٦) توزيع t   t Distribution

في معظم الأبحاث وغالبا يكون تباين المجتمع الذى تختار منه العينات مجهولا. للعينات العشوائية من الحجم $n \geq 30$ فإن التقدير الجيد للمعلمة σ^2 هو s^2. إذا كانت $n \geq 30$

واستبدلنا σ بالقيمة S في صيغة Z لنظرية (٧–٥) فإن :

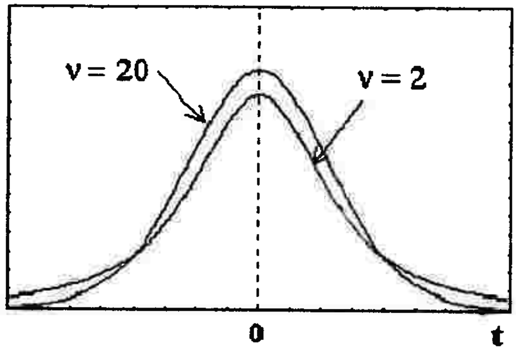
$$z = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

هي قيمة لمتغير عشوائى Z تقريبا يتبع التوزيع الطبيعى القياسى . أما إذا كان حجم العينة صغير ($n < 30$) فإن قيم $(\bar{x} - \mu)/(s/\sqrt{n})$ لا تتبع التوزيع الطبيعى القياسى . في هذه الحالة يكون اهتمامنا بتوزيع لإحصاء ما سوف نرمز له بالرمز T ، والذى قيمه تعطى من الصيغة التالية :

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

لقد تمكن ستيودنت "Student" وهو لقب لعالم إحصائي، كان ينشر أبحاثه بتوقيع ستيودنت، أن يشتق العبارة المضبوطة لتوزيع t ويسمى هذا التوزيع في كتب الإحصاء المختلفة " توزيع t" أو " توزيع ت ". يشبه توزيع t التوزيع الطبيعى القياسى فكلاهما متماثل حول الصفر كما أن كلا التوزيعيين لهما شكل الناقوس ولكن توزيع t أكثر تشتتا وذلك راجع إلى الحقيقة أن قيم t تعتمد على الاختلاف في قيمتي $\bar{x}$ و s^2 بينما قيم z تعتمد فقط على التغير في قيمة $\bar{x}$ من عينة إلى أخرى. يختلف توزيع المتغيرT عن المتغير Z في إن التباين يعتمد على حجم العينة n ودائما أكبر من الواحد الصحيح ، فقط عندما $n \to \infty$ فإن التوزيعين يتساويان. المقام (n-1) والذي يظهر في صيغه s^2 يسمى درجات الحرية degree of freedom المرتبط بتباين العينة s^2. بتكرار المعاينة من الحجم n وحساب $\bar{x}$ و s^2 لكل عينة، فإن قيم t المقابلة يقال أنها تتبع توزيع t بدرجات حرية ν ، حيث $\nu = n-1$. وعلى ذلك سوف يكون لدينا منحنيات t مختلفة أو توزيع t لكل حجم عينه. من خصائص توزيع t أنه كلما كبرت درجات الحرية ν زاد ارتفاع منحنى t وأصبح أكثر تدببا أي اقل تشتتا وفي النهاية ينطبق على منحنى التوزيع الطبيعي القياسي.

المنحنى في شكل (٧ – ١٤) بدرجات حرية $\nu = 2$ يمثل توزيع كل قيم t المحسوبة من عينات عشوائية من الحجم $n = 3$ تكرر اختيارها من مجتمع طبيعي. بنفس الشكل، المنحنى بدرجات حرية $\nu = 20$ يمثل توزيع كل قيم t المحسوبة من عينات من الحجم $n = 21$.



شكل (٧ –١٤)

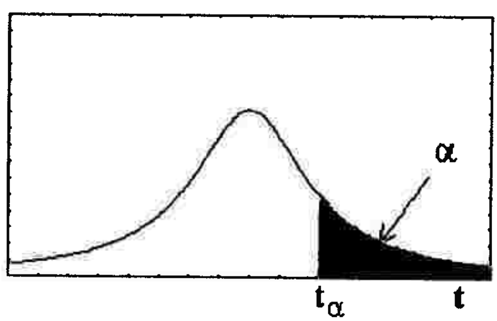
نظرية (٧–٩) إذا كان $\bar{x}$ و s^2 هما المتوسط الحسابي والتباين على التوالي لعينة عشوائية من الحجم n مأخوذة من مجتمع طبيعي له متوسط μ وتباين σ^2 غير معروف فإن :

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

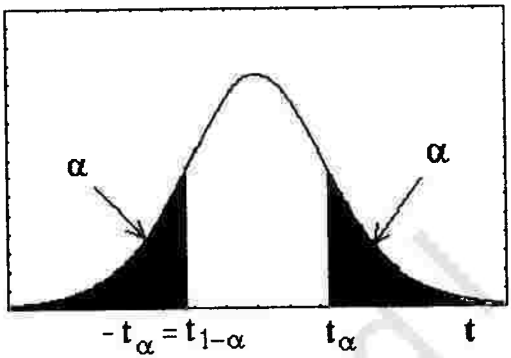
هي قيمة لمتغير عشوائي T له توزيع t بدرجات حرية $v = n-1$.

بفرض أن t_α ترمز لقيمة t التي توجد على المحور الأفقي تحت منحنى توزيع t بدرجات حرية v والتي المساحة على يمينها قدرها α كما هو موضح في شكل (٧–١٥).

شكل (٧–١٥)



الجدول في ملحق (٤) يعطى قيم t_α التي تناظر الاحتمال α لدرجات حرية ν حيث α تأخذ القيم التالية : .0005, .001, .005, .01, .025, .05, .10. ودرجات الحرية تأخذ القيم من $\nu = 1$ إلى $\nu = \infty$. يوضح الصف الثاني من الجدول قيم α والعمود الأول من الشمال قيم درجات الحرية ν. أما محتويات الجدول فهى القيم t_α . ولأن المنحنى متماثل فإن $t_{1-\alpha} = -t_\alpha$ كما هو موضح في شكل (٧-١٦).



شكل (٧-١٦)

مثال (٧-١٥) أوجد (أ) قيمة $t_{.005}$, $\nu = 15$.

(ب) قيمة $t_{.995}$, $\nu = 15$.

الحل .

(أ) بالبحث في جدول توزيع t في ملحق (٤) عند تقاطع الصف $\nu = 15$ والعمود $\alpha = .005$ نجد أن $t_{.005} = 2.947$.

(ب) باستخدام خاصية التماثل لمنحنى توزيع t فإن

$t_{.995} = -t_{.005}$ أى أن $t_{.995} = -2.947$.

مثال (٧-١٦) أوجد قيمة α حيث

$$\nu = 16 \quad , t_\alpha = -1.746$$

الحل .

حيث أن قيمة t سالبة فإنها تقع في الذيل الأيسر من توزيع t وباستخدام خاصية التماثل لمنحنى توزيع t فإن :

$$t_{1-\alpha} = -t_{\alpha} = 1.746$$

ومن جدول توزيع t في ملحق (٤) فإن $1-\alpha = .05$ ومنها $\alpha = .95$.

مثال (٧ –١٧) إذا كانت أعمار المصابيح الكهربائية المنتجة بواسطة أحد المصانع تتبع توزيعاً طبيعياً ، ويدعي صاحب المصنع أن متوسط أعمار هذه المصابيح هو $\mu = 500$ ساعة. لمتابعة جودة الإنتاج يأخذ 20 مصباحا كل شهر ، ويحقق الإنتاج للمواصفات القياسية إذا كانت قيمة t المحسوبة من عينة عشوائية من الحجم $n = 20$ تقع في الفترة($-t_{.05}$, $t_{.05}$) ما الاستنتاج الذي يمكن وضعه عند اختيار عينة عشوائية من الحجم $n = 20$ لها متوسط حسابي $\bar{x} = 530$ وانحراف معياري $s = 20$ ؟

الحل .

من جدول توزيع t في ملحق (٤) نجد أن $t_{.05} = 1.729$ عند درجات حرية $\nu = 19$.

وعلى ذلك فإن المنتج يحقق المواصفات القياسية إذا كانت قيمة t المحسوبة من عينة عشوائية من الحجم $n = 20$ مصباحا تقع في الفترة (-1.729 , 1.729) . ونجد أن :

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{530 - 500}{\frac{20}{\sqrt{20}}} = 6.708.$$

لا تقع في الفترة (-1.729 , 1.729) إذا كانت $\mu > 500$ فإن قيمة t المحسوبة من العينة تكون جيدة وتعني أن الإنتاج أفضل من المطلوب.

إذا كانت لدينا عينتان عشوائيتان مأخوذتان من مجتمعين طبيعيين بمتوسطي μ_1, μ_2 وسوف تكون نظرية (٧–٦) مفيدة فقط إذا كانت العينتان مستقلتان وتبايني المجتمعين σ_1^2, σ_2^2 معلومتان أو يمكن تقديرهما من عينات عشوائية من الحجم n_1, n_2 حيث $n_1 \geq 30$, $n_2 \geq 30$.

إذا كانت σ_1^2, σ_2^2 مجهولتان كما يحدث في معظم الحالات ، فإن التوزيع المضبوط للمتغير Z في نظرية (٧–٦) لا يكون معروفا فيما عدا لو فرضنا أن $\sigma_1^2 = \sigma_2^2 = \sigma^2$ والتباين يمكن تقديره من الصيغة التالية :

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

وعلى ذلك يكون :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

هي قيمة للإحصاء T الذى يخضع لتوزيع t بدرجات حرية $\nu = n_1 + n_2 - 2$.

نظرية (٧–١٠) إذا كان $s_1^2, \bar{x}_1$ يمثلان المتوسط والتباين على التوالي لعينة عشوائية من الحجم n_1 مأخوذة من مجتمع طبيعى بمتوسط μ_1 وتباين مجهول σ_1^2 وإذا كانت $s_2^2, \bar{x}_2$ يمثلان المتوسط والتباين على التوالي لعينة عشوائية من الحجم n_2 مأخوذة من مجتمع طبيعى بمتوسط μ_2 وتباين σ_2^2 مجهول وإذا كانت n_1 مستقلة عن n_2 و $\sigma_1^2 = \sigma_2^2$ فإن :

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

هي قيمة لمتغير عشوائى T له توزيع t بدرجات حرية $\nu = n_1 + n_2 - 2$.

مثال (٧–١٨) يقوم مصنع بإنتاج نوعين من المصابيح الكهربائية A,B. المصابيح من النوع A لها متوسط عمر أطول 100 ساعة عن متوسط عمر المصابيح من النوع B. التباين لكلا النوعين واحد. يختار شهريا 15 مصباحا من النوع الأول، 10 مصابيح من النوع الثاني للاختبار وتحسب قيمة t. تحقق القيمة المواصفات القياسية إذا وقعت في الفترة $(-t_{0.01}\ ,\ t_{0.01})$ فإذا تم سحب في شهر ما عينة عشوائية من 15 مصباحا من المصنع A ووجد أن $s_1 = 50, \bar{x}_1 = 520$. أيضا تم سحب عينة عشوائية من المصنع B من الحجم $n_2=10$ ووجد أن $s_2 = 40\ ,\ \bar{x}_2 = 500$ هل الإنتاج يحقق المواصفات القياسية ؟

الحل . باستخدام جدول توزيع t في ملحق (٤) فإن $t_{0.01} = 2.5$ بدرجات حرية $\nu = 15 + 10 - 2 = 23$. وعلى ذلك الإنتاج يحقق المواصفات القياسية إذا كانت قيمة t المحسوبة تقع في الفترة , $(-2.5\ ,\ 2.5)$. التباين المتجمع s_p^2 هو :

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$= \frac{(14)(50)^2 + (9)(40)^2}{15 + 10 - 2} = 2147.826.$$

وبأخذ الجذر التربيعي للتباين المتجمع s_p^2 فإن $s_p = 46.3446$.وعلى ذلك :

$$t=\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

$$=\frac{(520-500)-(100)}{46.3446\sqrt{\frac{1}{15}+\frac{1}{10}}}=-4.228.$$

وبما أن t لا تقع في الفترة (2.5 , -2.5) فإن الإنتاج لا يحقق المواصفات القياسية .

في بعض الأحيان بدلا من استخدام طريقة العينات المستقلة فإنه غالبا ما تستخدم طريقة العينات المتزاوجة paired samples. ففي تجارب تغذية الحيوان عند مقارنة عليقتين، حيث توضع الحيوانات المتجانسة في أزواج ويشترط أن تكون هذه الأزواج على درجة عالية من التماثل وقد تختلف الأزواج فيما بينها إلا أن أفراد كل زوج تكون متماثلة ويعطى أحد أفراد كل زوج عليقة بينما يعطى الأخر العليقة الأخرى وبالتالى فإن المقارنة بين العليقتين تتم داخل مجموعات متجانسة. في بعض الأحيان يتم ازدواج المشاهدات لمفردات العينة نفسها. فمثلا لمعرفة تأثير دواء على ارتفاع ضغط الدم نختار عينة عشوائية من الحجم n من الأشخاص ويتم قياس ضغط الدم الخاص بهم في أول فترة زمنية ثم يعالجون بهذا الدواء وبعد فترة زمنية معينة يتم قياس ضغط الدم لهم مرة أخرى . أزواج المشاهدات سوف تكون (x_1 , y_1) , $(x_2 , y_2),\ldots,(x_n, y_n)$. الفروق لأزواج المشاهدات سوف تكون $d_i = x_i - y_i$, $i= 1, 2, \ldots, n$. هذه الفروق تمثل قيم لمتغير عشوائى D. متوسط مجتمع الفروق μ_D سوف يساوى الصفر إذا كان الدواء ليس له تأثير. وسوف نرمز لمتوسط الفروق في عينتنا بالرمز $\bar{d}$، وسوف تختلف $\bar{d}$ من عينة إلى أخرى ولذلك يعتبر $\bar{d}$ قيمة للإحصاء $\bar{D}$. كما يحسب تباين الفروق s_d^2 والذي يعتبر قيمة للإحصاء S_d^2 لأنه يتغير من عينة إلى أخرى.

نظرية (٧ – ١١) إذا كان $d_1, d_2, \ldots, d_n$ تمثل الفروق لعدد n من أزواج المشاهدات وإذا كانت الفروق التى عددها n تمثل عينة عشوائية لها متوسط $\bar{d}$ وتباين s_d^2 مأخوذة من مجتمع الفروق الطبيعي والذى له متوسط μ_D وتباين σ_D^2 فإن :

$$t=\frac{\bar{d}-\mu_D}{\frac{s_d}{\sqrt{n}}}$$

هي قيمة لمتغير عشوائى T له توزيع t بدرجات حرية $\nu = n-1$.

مثال (٧ –١٩) إذا كان من المعتقد أن أكل السمك يساعد على زيادة الذكاء . أجريت تجربة على 11 شخصا تم اختيارهم عشوائيا وأجرى لهم أحد اختبارات الذكاء ثم أعطى لهم طعام يحتوى أساسا على السمك وبعد فترة معينة أجرى لهم اختبار الذكاء مرة أخرى فكانت النتائج كالتالى :

قبل أكل السمك	96	109	104	120	120	100	80	111	90	102	103
بعد أكل السمك	97	112	105	105	117	101	89	114	105	105	109

بفرض أن مستوى الذكاء قبل وبعد أكل السمك يتبع توزيعا طبيعيا وإذا اعتبرنا أن السمك له تأثير على مستوى ذكاء الأشخاص إذا لم تقع قيمة t المحسوبة من العينة في الفترة $(-t_{.05}\ ,\ t_{.05})$ بدرجات حرية $\nu = n-1 = 9$ ، ما الاستنتاج الذي يمكن الحصول عليه من بيانات العينة السابقة ؟

الحل .

من جدول t في ملحق (٤) فإن $t_{.05}= 1.812$ بدرجات حرية $\nu = 10$. وعلى ذلك يعتبر أكل السمك له تأثير على مستوى ذكاء الأشخاص إذا وقعت قيمة t خارج الفترة $(-1.812\ ,\ 1.812)$. من بيانات العينة فإن الفروق :

-1, -3, -1, 15, 3, -1, -9, -3, -15, -3, -6

ومنها :

$$n = 11, \Sigma d_i = -24, \Sigma d_i^2 = 606$$

$$\bar{d} = \frac{\Sigma d_i}{n} = \frac{-24}{11} = -2.1818,$$

$$s_{\bar{d}} = \sqrt{\frac{1}{n-1}\left[\Sigma d_i^2 - \frac{(\Sigma d_i)^2}{n}\right]}$$

$$= \sqrt{\frac{1}{10}\left[606 - \frac{(-24)^2}{11}\right]} = 7.44067.$$

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} = \frac{2.1818}{7.44067/\sqrt{11}} = 0.97252.$$

وحيث أن قيمة t المحسوبة تقع فى الفترة (1.812 , 1.812-) فيمكن القول أن السمك ليس له تأثير على مستوى الذكاء .

(٧-٧) توزيع مربع كاى Chi - Square Distribution

إذا تكرر سحب عينات من الحجم n من توزيع طبيعى تباينه σ^2 وإذا تم حساب تباين العينة s^2 لكل عينة فإننا نحصل على قيم للإحصاء S^2 . التوزيع العينى للإحصاء S^2 له تطبيقات قليلة في الإحصاء . اهتمامنا سوف يكون في توزيع المتغير X^2 والتي تحسب قيمته من الصيغة الآتية :

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \quad .$$

توزيع المتغير العشوائى X^2 يسمى توزيع χ^2 (توزيع مربع كاى) بدرجات حرية $v = n-1$. كما ذكرنا سابقا فإن v تساوى المقام في صيغة s^2.

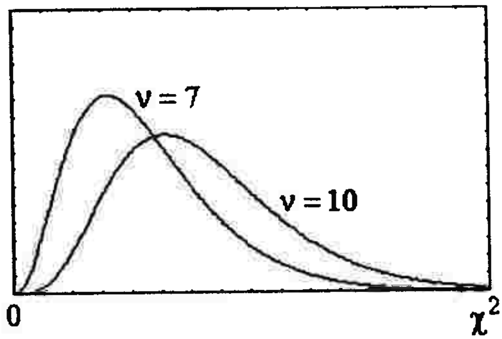
من الواضح أن قيم χ^2 لا يمكن أن تكون سالبة وعلى ذلك فإن منحنى توزيع χ^2 لا يمكن أن يكون متماثل حول الصفر. التوزيع العيني للإحصاء X^2 يمكن الحصول عليه باختيار عينات عشوائية متكررة من الحجم n من مجتمع طبيعى وحساب القيم χ^2 لكل عينه ثم يمكن الحصول على منحنى χ^2 بتمهيد المدرج التكرارى لقيم χ^2 ، يعتمد شكل المنحنى على قيم v. يوضح شكل (٧-١٧) منحنيان لتوزيع χ^2 بدرجات حرية $v = 7$ و $v = 10$ حيث يمثل المنحنى بدرجات حرية $v = 7$ توزيع قيم χ^2 المحسوبة من كل العينات من الحجم $n = 8$ من مجتمع طبيعى تباينه σ^2 . بنفس الشكل يمثل المنحنى بدرجات حرية $v = 10$ توزيع كل قيم χ^2 المحسوب من كل العينات من الحجم n=11.

نظرية (٧-١٢) إذا كان s^2 يمثل تباين العينة من الحجم n المأخوذة من مجتمع طبيعى له تباين σ^2 فإن :

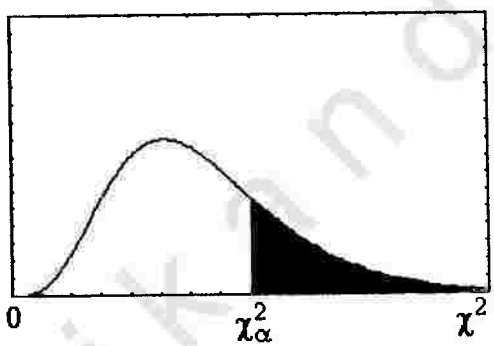
$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

هي قيمة لمتغير عشوائى X^2 له توزيع χ^2 بدرجات حرية $v = n-1$.

شكل (٧-١٧)



بفرض أن χ_α^2 ترمز لقيمة χ^2 التي توجد على المحور الأفقي تحت منحنى χ^2 بدرجات حرية v والتي تكون المساحة على يمينها قدرها α كما هو موضح في شكل (٧-١٨).



شكل (٧-١٨)

الجدول في ملحق (٥) يعطى قيم χ_α^2 وذلك لقيم مختلفة من α و v حيث α تأخذ القيم :

.995 , .99, .975, .95, .90, .10, .05, 0.025, .01, .005

ودرجات حرية من $v = 1$ إلى $v = 40$. يوضح الصف الثاني من الجدول قيم α والعمود الأول من الشمال قيم درجات الحرية أما محتويات الجدول فهي لقيم χ_α^2. وعلى ذلك للحصول على قيمة χ_α^2 بدرجات حرية $v = 6$ والتي تكون المساحة على يمينها تساوى 05. فإننا نبحث في الجدول عند تقاطع الصف الذى به $v = 6$ مع العمود 05. وعلى ذلك $\chi_{.05}^2 = 12.592$. ولعدم تماثل منحنى توزيع χ^2 فلا بد من

استخدام الجدول لإيجاد $\chi^2_{.95} = 1.635$ عند $v = 6$.

مثال (٧ – ٢٠) أوجد قيمة χ^2 لتوزيع χ^2 بدرجات حرية $v = 14$ والتي تكون المساحة على يمينها تساوى 0.01 .

الحل .

بالبحث في جدول توزيع χ^2 في ملحق (٥) عند تقاطع الصف $v = 14$ مع العمود $\alpha = 0.01$ نجد أن $\chi^2 = 29.141$.

مثال (٧–٢١) أوجد قيمة χ^2 التي تكون المساحة على يسارها تساوى 0.99 لتوزيع χ^2 بدرجات حرية $v = 4$.

الحل .

قيمة χ^2 التي تكون المساحة على يسارها تساوي .99 هي χ والتي تكون المساحة على يمينها تساوى .01 = .99 -1 وعلى ذلك فإن قيمة $\chi^2_{.01}$ بدرجات حرية $v = 4$ هي تلك القيمة في جدول توزيع χ^2 التي تقع عند تقاطع الصف $v = 4$ والعمود $\alpha = .01$ هى $\chi^2_{.01} = 13.277$.

مثال (٧–٢٢) أوجد قيمتي χ^2 لتوزيع χ^2 بدرجات حرية $v = 15$ اللتين تحصران بينهما 99% من المساحة تحت المنحنى بحيث أن المساحة في الطرف الأيسر للتوزيع تساوى المساحة في الطرف الأيمن للتوزيع.

الحل . المطلوب هنا هو إيجاد قيمتي χ^2 اللتين تقسمان المنحنى بحيث أن المساحة في الطرف الأيمن هى .005. وعلى ذلك قيمة χ^2 في الطرف الأيمن هى $\chi^2_{.005} = 32.799$. وللحصول على قيمة χ^2 في الطرف الأيسر والتي المساحة التي تقع على يسارها هى 0.005 وبالتالي فإن المساحة التى تقع على يمينها هى .995 = 1-.005 وعلى ذلك فإن قيمة χ^2 في الطرف الأيسر هى $\chi^2_{.995} = 4.600$ وبالتالي فإن القيمتين هما , 4.600 . 32.799

(٧ –٨) توزيع F   F Distribution

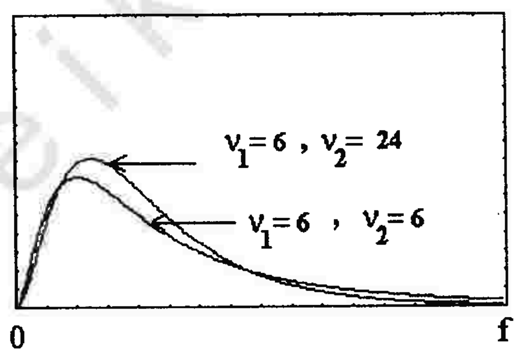
يعتبر توزيع F من التوزيعات الاحتمالية الهامة التى تستخدم في مجال الإحصاء التطبيقى . نظريا يمكن تعريف توزيع F (توزيع ف) كنسبة لتوزيعين مستقلين يتبعان توزيع χ^2 وكل منهما له درجات حرية خاصة به . فـإذا كانت f قيمة للمتغير العشوائى F ، فإن:

وكل منهما له درجات حرية خاصة به . فإذا كانت f قيمة للمتغير العشوائى F ، فإن:

$$f = \frac{\chi_1^2 / \nu_1}{\chi_2^2 / \nu_2} = \frac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2} = \frac{\sigma_2^2 s_1^2}{\sigma_1^2 s_2^2}.$$

حيث χ_1^2 هي قيمة لتوزيع χ^2 بدرجات حرية $\nu_1 = n_1 - 1$ و χ_2^2 هي قيمة لتوزيع χ^2 بدرجات حرية $\nu_2 = n_2 - 1$.

لحساب قيمة f نختار عينة عشوائية من الحجم n_1 من مجتمع طبيعى له تباين σ_1^2 ونحسب s_1^2 / σ_1^2 . أيضا نختار عينة عشوائية مستقلة من الحجم n_2 من مجتمع طبيعى آخر له تباين σ_2^2 ونحسب s_2^2 / σ_2^2. النسبة للقيمتين s_1^2 / σ_1^2 و s_2^2 / σ_2^2 تنتج قيمة f • توزيع كل القيمة الممكنة من f حيث s_1^2 / σ_1^2 يمثل البسط و s_2^2 / σ_2^2 يمثل المقام يسمى توزيع F بدرجات حرية ν_1, ν_2. إذا اعتبرنا كل النسب الممكنة حيث s_2^2 / σ_2^2 يمثل البسط و s_1^2 / σ_1^2 يمثل المقام ، في هذه الحالة نحصل على توزيع كل القيم الممكنة التى تتبع توزيع F ولكن بدرجات حرية ν_2, ν_1 . درجات الحرية المرتبطة بتباين القيمة التى في البسط دائما يوضع أولا متبوعا بدرجات الحرية المرتبطة بتباين العينة التى في المقام. وعلى ذلك منحنى توزيع F يعتمد ليس فقط على المعلمتين ν_1, ν_2. ولكن أيضا على ترتيبهما وبمجرد الحصول على القيمتين يمكن تعريف المنحنى • منحنيان لتوزيعF موضحة في شكل (٧-١٩)•

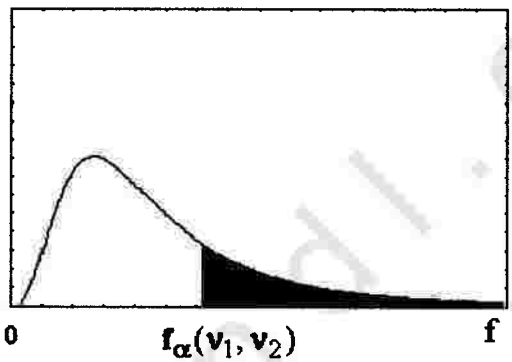


شكل (٧-١٩)

نظرية (٧-١٣) إذا كانت s_1^2 , s_2^2 تمثلان تبايني عينتين عشوائيتين مستقلتين من الحجم n_1 , n_2 مأخوذتين مجتمعين من طبيعيين بتباينى σ_1^2 , σ_2^2 على التوالي فإن :

$$f = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} = \frac{\sigma_2^2 s_1^2}{\sigma_1^2 s_2^2}.$$

هي قيمة لمتغير عشوائي F يتبـع توزيـع F بدرجـات حريـة ν_1, ν_2 • بفـرض أن $f_\alpha(\nu_1, \nu_2)$ ترمز لقيمة f على المحور الأفقي تحت منحنى توزيع F بدرجات حرية $\nu_1 = n_1 - 1$ و $\nu_2 = n_2 - 1$ والتي تكون المساحة على يمينها تساوى α والموضحة في شكل (٧-٢٠).



شكل (٧-٢٠)

لاستخراج قيم $f_\alpha(\nu_1, \nu_2)$ يوجد جدولان في ملحق (٦) وملحق (٧) ، الأول عند $\alpha=$ .05 والآخر عند $\alpha = .01$ وفي كل منهما يكون الصف الأول لقيم ν_1 والعمـــود الأول لقيم ν_2 أما محتويات الجدول فهو لقيم $f_\alpha(\nu_1, \nu_2)$. على سبيل المثال مـن جـدول توزيع F نلاحظ أن :

$$f_{.01}(5,7) = 7.46 \ , \ f_{.05}(1,4) = 7.71$$

$$f_{.01}(9,10) = 4.94 \ , \ f_{.05}(4,1) = 224.6$$

وباستخدام النظرية التالية يمكن استخدام جدول توزيع F في إيجاد $f_{1-\alpha}(\nu_1, \nu_2)$.

نظرية (٧-١٤) يمكن كتابة $f_{1-\alpha}(\nu_1, \nu_2)$ بدرجات حرية ν_1, ν_2 على الشكل :

$$f_{1-\alpha}(\nu_1,\nu_2)=\frac{1}{f_{\alpha}(\nu_2,\nu_1)}$$

وعلى ذلك فإن قيمة $f_{.95}(7,12)$ هى :

$$f_{.95}(7,12)=\frac{1}{f_{.05}(12,7)}=\frac{1}{3.57}=0.2801$$

حيث أن $f_{.05}(12\ ,\ 7)$ مستخرجة من جدول توزيع F في ملحق (٦) عند مستوى معنوية $\alpha = .05$ ودرجات حرية $\nu_1 = 12,\ \nu_2 = 7$ ،

<u>تماريـــن :</u>

١- مجتمع محدود يتكون من القيم 2, 4, 6 (أ) أوجد المـــدرج التكـــرارى للتوزيـــع العيني للإحصاء $\overline{X}$ عند سحب عينات عشوائية من الحجم $n = 4$ مع الإرجـــاع (ب) تحقق أن $\mu_{\overline{X}} = \mu$ و $\sigma^2_{\overline{X}} = \dfrac{\sigma^2}{n}$.

٢- يتكون مجتمع من القيم 1, 1, 1, 2, 2, 3, 4 (أ) أوجد كل العينات الممكنة من الحجم $n = 2$ والتي يمكن اختيارها من هذا المجتمع بدون إرجاع (ب) أثبت أن :

$\mu_{\overline{X}} = \mu$ و $\sigma^2_{\overline{X}} = \dfrac{\sigma^2}{n}\dfrac{N-n}{N-1}$.

٣- إذا كانت 3, 4, 7, 9, 12 تمثل مفردات المجتمع محل الدراسة فإذا سحبت عينـــة مكونه من مفردتين :

(أ) أحسب التوقع والتباين للمجتمع .

(ب) حدد العينات الممكن سحبها بحيث يكون حجم كل عينة مفردتين إذا كان السحب مع الإرجاع .

(ج) حدد العينات الممكن سحبها إذا كان السحب بدون إرجاع .

(د) أوجد التوقع والتباين للعينات المسحوبة ف (ب) و (ج) .

٤- كم عينة عشوائية مختلفة حجمها $n = 3$ يمكن اختيارها بإحلال ثم بدون إحلال من مجتمعات محدودة مكونة من $N = 25\ ,\ N = 20\ ,\ N = 15$ ثم أوجـــد احتمـــال

اختيار كل عينة .

-٥- كم عينة عشوائية مختلفة حجمها $n = 2$ يمكن اختيارها بإحلال ثم بدون إحلال من مجتمعات محدودة مكونة من $N = 10$, $N = 15$, $N = 30$ ثم أوجد احتمال اختيار كل عينة .

-٦- ما هي قيمة معامل التصحيح للمجتمع المحدود عندما تكون :

(أ) $N = 100$, $n = 10$ (ب) $N = 200$, $n = 5$

(ج) $N = 50$, $n = 3$ (د) $N = 300$, $n = 15$

-٧- يقوم مصنع بإنتاج مصابيح كهربائية بمتوسط عمر 1900 ساعة وانحراف معياري 200 ساعة. احسب احتمال أن يكون متوسط العينة المبنى على عينة حجمها 100 مصباح أكبر من 1850 وأقل من 1920.

-٨- صممت آلة للشراب الرطب بحيث أن كمية الشراب الذى تخرجه يتبع التوزيع الطبيعى بمتوسط $\mu = 8$ أوقية للكوب وانحراف معياري 0.5 أوقية. تختبر الآلة دوريا بأخذ عينة من 9 أكواب وحساب متوسط العينة. إذا كان متوسط كمية الشراب المحسوب من 9 أكواب تقع في الفترة $\mu_{\bar{X}} \pm 2\sigma_{\bar{X}}$ فالآلة تعمل بطريقة صحيحة و غير ذلك لا بد من فحص الآلة واتخاذ اللازم . ما هو القرار الذى يجب اتخاذه عند سحب عينة عشوائية من 9 أكواب من الشراب الرطب وكان متوسط كمية الشراب للكوب 8.4 أوقية ؟

-٩- إذا كان ضغط الدم لمجموعة من الأفراد متوسطة 85 وانحرافه المعياري 5. أخذنا عينة عشوائية حجمها 40 فردا من هذه المجموعة، أوجد احتمال أن متوسط ضغط الدم في العينة أكبر من 80.

-١٠- إذا كان متوسط الدخل الشهري للأسر في إحدى المدن هو 7000 ريال بانحراف معياري 500 ريال. اختيرت عينة عشوائية حجمها 100 أسرة من هذه المدينة أوجد احتمال :

(أ) أن يقل متوسط دخل الأسرة في العينة عن 4800 ريال

(ب) أن يتراوح متوسط دخل الأسرة في العينة بين 6000 , 6500.

-١١- صممت إحدى الشركات سيارة بحيث أن أكبر حمولة لها 3000 كم ، وتتسع إلى 30 راكبا . إذا علمت أن أوزان الأشخاص الذين يستعملون هذه السيارة تخضع لتوزيع طبيعي متوسطة 70 كم وانحرافه المعياري 10 كم. أحسب احتمال أن تتحمل اكثر من طاقتها إذا كان مجموع الأشخاص الذين يستعملونها أكبر من 3000 كجم.

-١٢- إذا علمت أن درجات الطلاب في مقرر الإحصاء تخضع لتوزيع طبيعي وسطه 70 وانحرافه المعياري 9. أخذت عينة عشوائية حجمها 15 طالبا. أحسب احتمال أن يزيد متوسط علامات العينة عن 74.

-١٣- إذا كان X متغير عشوائيا وسطه 180.5 وانحرافه المعياري 12. من هذا التوزيع أخذت عينه حجمها 36. ما هو احتمال أن يزيد وسط العينة عن 190 ؟

-١٤- إذا كان X متغيرا عشوائيا يمثل أوزان أكياس الطحينة التى تنتجها إحدى المؤسسات وكان X خاضعا لتوزيع طبيعى وسطه 50 كيلو جراما وانحرافه المعيارى 5 كيلو جرام. أخذت عينة حجمها 9 أكياس من إنتاج هذه المؤسسة أحسب احتمال أن يزيد الوسط الحسابى لهذه العينة عن 51 كيلو جرام.

-١٥- إذا كان عدد الحوادث الأسبوعية على إحدى الطرق يتبع توزيع بواسون بمتوسط 5 حوادث فإذا أخذت عينة من 60 أسبوعا فما هو احتمال أن يكون متوسط عدد الحوادث فيها أقل من 3 حادث ؟

-١٦- إذا كان متوسط أعمار مرض السكر هو 60 سنة انحراف معياري 15 سنة. اختيرت عينة عشوائية حجمها 30 مريضا بالسكر. أوجد احتمال :

أ- أن يقل متوسط العمر في العينة عن 50 سنة .

ب - أن يتراوح متوسط العمر في العينة بين 40 ,50 سنة .

-١٧- سحبت عينة عشوائية من الحجم 36 من مجتمع كبير متوسطة $\mu = 80$ وانحرافه المعياري 18 فما هى النسبة المئوية من متوسطات العينات التي :

أ- تقع بين 71 و 89 . ب- أكبر من 89

ج- تقع بين 74 و 86 د- أقل من 74

-١٨- يرغب باحث في تقدير متوسط مجتمع وذلك بالاعتماد على عينة عشوائية مسحوبة من هذا المجتمع وتكون كبيرة بدرجة كافية وذلك تحت شرط :

$$P\left|\overline{X}-\mu\right|<\sigma/4=0.95$$

ما هي أقل حجم للعينة يجب على الباحث استخدامه ؟

-١٩- إذا سحبت كل العينات الممكنة من الحجم 25 من مجتمع طبيعي بمتوسط 50 وانحراف معياري 5. ما هو الاحتمال أن متوسط العينة $\overline{X}$ سوف يقع في الفترة :

$$(\mu_{\overline{X}}-1.96\ \ \sigma_{\overline{X}},\mu_{\overline{X}}+1.96\ \ \sigma_{\overline{X}})$$

-٢٠- إذا كانت أطوال 1000 طالب تقريبا تتبع التوزيع الطبيعي بمتوسط 68.5 بوصة وانحراف معياري 2.7 بوصة إذا سحبنا 200 عينة عشوائية من الحجم n =25 من هذا المجتمع وثم حساب المتوسطات . المطلوب :

أ- المتوسط والانحراف المعيارى للتوزيع العينى للمتوسط .

ب- عدد متوسطات العينات التى تقع في الفترة (68 , 69.2) .

ج- عدد متوسطات العينات التى تقل عن 67.2 .

-٢١- فتائل حريرية متوسط نقطة قطعها $\mu=23$ رطلا بانحراف معيارى $\sigma=0.4$ رطلا . اختيرت عينة عشوائية حجمها 40 فتيلة وذلك لإيجاد نقطة قطعها . ما هو احتمال أن متوسط نقطة القطع للعينة سيكون أكبر من 25.5 ؟

-٢٢- في دراسة عن تلوث الهواء بأكسيد الكبريت المنبعث من أحد المصانع كان متوسط التلوث $\mu=18$ طنا بانحراف معياري $\sigma=5.5$ طنا فإذا اختيرت عينة عشوائية مكونة من قراءات 90 يوما أحسب احتمال أن يكون متوسط العينة أكبر من 18.5 .

-٢٣- إذا كان X متغيرا عشوائيا يمثل أطوال الطلاب في جامعة ما ، وكان X خاضعا لتوزيع طبيعي متوسطة 160 سم وانحرافه المعيارى 20 سم . أخذت عينة عشوائية حجمها 15 طالبا من هذه الجامعة أحسب احتمال أن يقل متوسط العينة لأطوال أفراد هذه العينة عن 160 سم .

-٢٤- إذا كان $\overline{X}$ يمثل المتوسط لعينة من الحجم $n_1=2$ مع الإرجاع من مجتمع محدود له القيم 3, 4, 8 بنفس الشكل إذا كان $\overline{X}_2$ يمثل متوسط عينة من الحجم $n_2=2$ مع الإرجاع من مجتمع 2, 2, 4

أ- أوجد المدرج التكرارى للتوزيع العينى للإحصاء $\overline{X}_1 - \overline{X}_2$

ب - تحقق من أن $\mu_{\overline{X}_1-\overline{X}_2} = \mu_1 - \mu_2$, $\sigma^2_{\overline{X}_1-\overline{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

-٢٥- سحبت عينة عشوائية من الحجم $n = 25$ من توزيع طبيعى لــه متوســط 80 وانحراف معيارى 5. أيضا أخذت عينة عشوائية أخرى من الحجم 36 من مجتمع آخــر طبيعى له متوسط 75 وانحراف معيارى 3. أوجد الاحتمال أن متوسط العينة المحســوبة من 25 مفردة يزيد عن متوسط العينة المحسوب من 36 مفردة على الأقل 3.4 .

-٢٦- عينات عشوائية من الحجم 100 سحبت بدون إرجاع من مجتمعــين A , B فإذا كان لدينا المعلومات التالية :

$$\mu_1 = 10 \ , \ \sigma_1 = 2 \ , \ \mu_2 = 8 \ , \ \sigma_2 = 1$$

أوجد أ- $E(\overline{X}_1 - \overline{X}_2)$ ب - $\sigma^2_{\overline{X}_1-\overline{X}_2}$

ج- الاحتمال أن الفرق $\overline{X}_1 - \overline{X}_2$ أقل من 1.5 .

د - الاحتمال أن الفرق $\overline{X}_1 - \overline{X}_2$ أكبر من 1.75 ولكن أقل من 2.5

-٢٧- أجرى بحث على أربعة طلبة لمعرفة هل يدخنون أم لا . فإذا كــانت نتــائج البحث كالآتى : No, yes, No, No

أ- أوجد نسبة المدخنين في المجتمع .

ب - أوجد التوزيع العينى للإحصاء $\hat{P}$ (نسبة المدخنــين) وذلــك في حالــة السحب بإرجاع أو بدون إرجاع ثم أوجد متوسط وتباين التوزيع إذا كان حجم العينــة التى تم اختيارها $n=2$.

-٢٨- أجرى بحث على 5 أفراد لمعرفة حامل ميكروب معين فكانت النتائج كما يلي حامل للميكروب ، غير حامل ، حامل للميكروب ، غير حامل ، غير حامل المطلوب :

أ- نسبة الحاملين للميكروب في المجتمع .

ب - إذا اختيرت نسبة عشوائية من فردين أوجد التوزيع العين للإحصاء $\hat{P}$ في

حالة السحب بإرجاع والسحب بدون إرجاع .

-٢٩- أجرى بحث لاستطلاع آراء 8 أفراد على منتج ما فكانت النتائج

yes , yes , yes , yes , yes , No, No, No المطلوب :

أ- نسبة الأفراد المفضلون للمنتج .

ب - إذا اختيرت عينة عشوائية من فردين أوجد التوزيع العيني للإحصاء $\hat{P}$ في حالة السحب بإرجاع والسحب بدون إرجاع .

-٣٠- إذا كانت نسبة المدخنين من الأفراد الذكور البالغين في إحدى المدن 20% فإذا اختيرت عينة عشوائية من 200 شخص أوجد احتمال أن يقل عدد المدخنين بينهم عن 3 أشخاص .

-٣١- في شركة كبيرة من 4000 موظف وجد أن الغياب في يوم الاثنين يمثل 8% لعدة سنوات. إذا اختيرت عينة عشوائية من 1500 موظف في يوم الاثنين. أوجد احتمال أن تكون نسبة الغياب أكبر من 0.07 .

-٣٢- إذا كان 45% من الأفراد من أصحاب الأملاك في بلد كبير يملكون سيارتين على الأقل. ما هو الاحتمال في عينة عشوائية من 150 شخص من هذا البلد تكون نسبة امتلاك سيارتين على الأقل 0.55 .

-٣٣- في مجتمع معين كانت نسبة المعاناة من مرض ما هو 0.2. أخذت عينة عشوائية من هذا المجتمع حجمها 250 (السحب بإرجاع) أوجد :

أ- احتمال أن تكون النسبة في هذه العينة أقل من 0.3 .

ب - بفرض أن السحب بدون إرجاع وحجم المجتمع 500 أوجد احتمال أن تكون النسبة في العينة أكبر من 0.4 .

-٣٤- في دولة ما معروف أن 70% من المرشحين في الانتخابات يناهضون قضية ما. فإذا اختيرت عينة عشوائية من هذه الدولة مكون من 100 مرشح أوجد احتمال أن تكون نسبة المرشحين الذين يناهضون القضية في العينة بين 80% , 60% .

-٣٥- إذا علم أن نسبة البيض التالف التى تنتجها أحد مراكز إنتاج البيض هى 0.2 . أشترى شخص 300 بيضة من إنتاج هذا المركز أوجد احتمال أن يجد من بينها 10 بيضة على الأقل تالفة .

-٣٦- إذا كان لدينا آلتين وكانت نسبة المعيب للآلة الأولى هو 15% والمعيب للآلة الثانية هو 12 % ، سحبت عينتان من إنتاج الآلتين حجمهما 650 ,600 وحدة على التوالي . فإذا كانت نسبة المعيب للعينة الأولى هو $\hat{P}_1$ ونسبة المعيب في العينة الثانية $\hat{P}_2$. احسب قيمة الاحتمال $P(\hat{P}_1 - \hat{P}_2 \le 0.2)$.

-٣٧- أوجد : أ - $t_{0.99}$, $\nu = 10$ ب - $t_{.05}$, $\nu = 15$

ج- $t_{.025}$, $\nu = 17$ د - $t_{.99}$, $\nu = 18$

-٣٨- أوجد قيمة t_α عندما $\nu = 23$ بحيث أن $P(-t_\alpha < T < t_\alpha) = .99$.

-٣٩- أوجد قيمة α التي تحقق الآتي : $t_\alpha = 1.33$ و $\nu = 18$.

-٤٠- أوجد الاحتمال أن المتغير العشوائى T والذى يتبع توزيع t :

أ- أكبر من 1.74 عند $\nu = 17$ ب - أقل من 1.323 عند $\nu = 21$

-٤١- يقوم مصنع بإنتاج مصابيح كهربائية متوسط عمرها 500 ساعة . للمحافظة على هذا الإنتاج يقوم المصنع شهريا باختيار 25 مصباحا. إذا وقعت قيمة t المحسوبة في الفترة $(t_{.05}\ ,\ -t_{05})$ فإن الإنتاج يكون محققا للمواصفات القياسية. ما هو الاستنتاج الذى يمكن اتخاذه عند اختيار عينة لها متوسط $\bar{x} = 518$ وانحراف معياري $s=40$ ساعة ؟ وذلك تحت فرض أن العينة تم اختيارها من مجتمع طبيعي .

-٤٢- أوجد المئين العاشر والمئين التسعين P_{90} لتوزيع t بدرجات حرية $\nu = 15$.

-٤٣- أوجد الربيع الأول Q_1 والربيع الثالث Q_3 والوسيط Q_2 لتوزيع t بدرجات حرية $\nu = 20$.

-٤٤- إذا كانت الأجور اليومية لعمال إحدى الشركات في مدينة ما تخضع للتوزيع الطبيعي بمتوسط $\mu_1 = 5$ وانحراف معياري $\sigma_1 = 0.5$. وكانت الأجور اليومية لعمال شركة مماثلة في مدينة أخرى تخضع للتوزيع الطبيعي بمتوسط $\mu_2 = 4$ وانحراف معياري $\sigma_1 = 10$ وبفرض أننا سحبنا عينة عشوائية من المجتمع الأول من الحجم $n_1 = 25$ وعينة أخرى من المجتمع الثانى من الحجم $n_2 = 20$ أوجد $P(\overline{X}_1 - \overline{X}_2 \le 0.75)$.

٤٥- اختيرت عينة عشوائية من مجتمع طبيعي له متوسط $\bar{x}_1 = 20$ وانحراف معياري $s_1 = 2$. كما اختيرت عينة أخرى عشوائية من الحجم $n_2 = 12$ من مجتمع آخر له متوسط $\bar{x}_2 = 24$ وانحراف معياري $s_2 = 6$ فإذا كانت $\mu_2 = 19$, $\mu_1 = 22$ و σ_1^2 , σ_2^2 مجهولتان ولكن تقريبا متساويتان أوجد قيمة t .

٤٦- طبق اختبار للقدرة على التفكير الناقد على مجموعة من المراهقين قبل حضورهم برنامج أعد لهذا الفرض مدته 40 أسبوعا وبعد حضورهم للبرنامج. فإذا اختيرت عينة عشوائية من الحجم $n = 10$ وكان متوسط الفروق $\bar{d} = 5$ بانحراف معياري $s_{\bar{d}} = 0.5$ أوجد قيمة t .

٤٧- أوجد النقاط التالية من جدول توزيع χ^2 في ملحق (٥) مع التوضيح بالرسم :

أ- $\chi^2_{.95}$, $\nu = 12$ ب- $\chi^2_{.05}$ و $\nu = 1$

٤٨- لتوزيع χ^2 أوجد

أ- $\chi^2_{.001}$ $\nu = 18,$ ب- $\chi^2_{.975}$ $\nu = 29,$

ج- χ^2_α بحيث أن $P(X^2 \le \chi^2_\alpha) = .99$ $\nu = 4,$.

٤٩- أوجد النقاط التالية من جدول توزيع F في ملحق (٦) مع التوضيح بالرسم $f_{.95}(6, 10)$, $f_{.05}(12,7)$.

٥٠- أوجد الاحتمال أن المتغير العشوائي F أكبر من 2.85 إذا كانت F له توزيع F بدرجات حرية $\nu_1 = 8, \nu_2 = 12$.

٥١- أوجد القيمة $f_\alpha(12,8)$ إذا كانت $P(F < f_\alpha(12,8)) = 0.05$.

٥٢- أوجد القيمة α لكل من القيم التالية مع التوضيح بالرسم $f_\alpha(9,8) = 3.39$, $f_\alpha(3,20) = 3.1$.

٥٣- أوجد الاحتمال أنه من عينة عشوائية من 25 مفردة من مجتمع طبيعي له تباين $\sigma^2 = 6$ سوف يكون التباين S^2

أ - أكبر من 9.1 ب- بين 4.62 , 10.745

٥٤- إذا كانت s_1^2, s_2^2 تمثلان تباين عينتين عشوائيتين من الحجم $n_1 = 31$, $n_2 = 25$ تم اختيارهما من مجتمعين طبيعيين . بحيث $\sigma_1^2 = 10, \sigma_2^2 = 15$ أوجد :

$P(S_1^2 / S_2^2) > 1.26)$.

٥٥- إذا كانت s_1^2, s_2^2 تمثلان تباين عينتين عشوائيتين من الحجم $n_1 = 8$, $n_2 = 12$ مسحوبتين من مجتمعين طبيعيين تحت فرض أن $\sigma_1^2 = \sigma_2^2$. أوجد الاحتمال $P(S_1^2 / S_2^2 < 4.89)$.