السلايدات الهامة والمهمة للتحليل الاحصائي

تجميع: عباسكو1

خواص العمليات الجبرية لإتحاد الحوادث:

-إذا كانت A و B و C ثلاث حوادث فإن :

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

ويعني ذلك توزيع الإتحاد على التقاطع.

– وكذلك هناك خاصية التبديل والتي تعني أن

$$(A \cup B) = (B \cup A)$$

تابع العمليات على المجموعات :-

المكملة أو المتممة :-

يقال أنإذا كانت تحتوي على جميع عناصر A مكملة المجموعة $\bar{A}$ المجموعة الكلية U باستثناء عناصر A.

مثال

إذا كان U = { 1,2,3,4,5,6,7,8,9,10} وA={2,4,6,8,10} أوجد

الحل

$\bar{A}$ = {1,3,5,7,9}

تقاطع الحوادث Events Intersection :

خواص العمليات الجبرية لتقاطع الحوادث:

– إذا كانت A و B و C ثلاث حوادث فإن :

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

ويعني ذلك توزيع التقاطع على الإتحاد.

– وكذلك هناك خاصية التبديل والتي تعني أن $(A \cap B) = (B \cap A)$

أمثلة وتمارين

و = $\cap$
أو = U

يراد شراء ثلاثة أنواع من اللحوم من جزار معين، فإذا رمزنا للحم الدجاج بـ A ولحم الضأن بـ B ، ولحم العجل بـ C فإن :

– توفر أنواع اللحوم الثلاثة يعني توفر لحم A و B و C أي بمعنى : $A \cap B \cap C$

– عدم توفر أي نوع من اللحوم يعني عدم توفر A و B و C أو كلها أي بمعنى :

$$\overline{A} \cap \overline{B} \cap \overline{C}$$

– توفر نوع واحد من اللحوم على الأقل هو توفر A أو B أو C أو كلها أي بمعنى :

$$A \cup B \cup C$$

– توفر نوع A فقط يعني : $A \cap \overline{B} \cap \overline{C}$

– توفر نوع واحد من اللحم يعني إما توفر A وعدم توفر النوعين الآخرين أو توفر B وعدم توفر النوعين الآخرين ، أو توفر C وعدم توفر النوعين الآخرين أي بمعنى :

$$(A \cap \overline{B} \cap \overline{C}) \cup (\overline{A} \cap B \cap \overline{C}) \cup (\overline{A} \cap \overline{B} \cap C)$$

1-التجربة العشوائية Random Experiment :

هي تلك التجربة التي تكون جميع نتائجها معلومة مسبقا ولكن لا يمكن التنبؤ بحدوث أي من هذه النتائج بصفة مؤكدة مثلا:

رمي حجر نرد مرة واحدة يعتبر تجربة عشوائية، حيث نعلم جميع قيم نتائج التجربة وهي إما 1 أو 2 أو 3 أو 4 أو 5 أو 6 ولكن لا يمكن تحديد أي الأرقام يظهر إلى الأعلى بصورة مؤكدة قبل إجراء التجربة.

-رمي عملة معدنية مرة واحدة أو عدد من المرات يعتبر تجربة عشوائية معروف جميع نتائجها قبل أن تبدأ التجربة، ولكن لا يمكن الجزم بظهور أي منها في رمية معينة.

-المشاركة في سباق الخيل لحصان معين يعتبر تجربة عشوائية فهو إما أن يفوز أو يخسر أو يتعادل.

2-الحادث والفراغ العيني:

فراغ العينة هو المجموعة الشاملة التي تحتوي على جميع النتائج الممكنة للتجربة العشوائية ويرمز له بالرمز Ω ويطلق علية الحالات الممكنة .

افترض أننا نقوم بإجراء تجربة ما كرمي زهرة النرد مثلاً ونلاحظ كل النتائج الممكنة وهي ظهور أحد الأوجه الستة 1 أو 2 أو 3 أو 4 أو 5 أو 6 ونفترض أننا مهتمون بظهور رقم فردي أي 1 أو 3 أو 5 من التجربة . وهكذا فإن عملية رمي الزهرة تسمى **تجربة** (Experiment) وظهور رقم فردي هو محل اهتمامنا يسمى **حادثاً** (Event) ومجموعة جميع الحالات الممكنة الظهور تسمى **بالفراغ العيني** (Sample Space) ويلاحظ أن الحادث قد يكون حالة او اكثر من الفراغ العيني .

الحادثة هي مجموعة جزئية من فراغ العينة وتمثل مجموعة النتائج التي تحقق الحدث وتسمى أيضا الحالات المواتية Favorable Cases ، فمثلا الحصول على رقم زوجي في تجربة رمي زهرة النرد مرة واحدة تكون الحادثة هي {2 ، 4 ، 6} ، ويمكن أن تحتوي الحادثة على عنصر واحد أو أكثر.

3- أنواع الحوادث :

أ- الحوادث المتنافية (Mutually Exclusive Events)

يقال عن الحادثين A و B أنهما متنافيان إذا استحال حدوثهما معا. فمثلاً عند رمي عملة معدنية لا يمكن الحصول على وجهين في وقت واحد.

ب-الحوادث المستقلة (Independent Events)

يعتبر الحادثين A أو B حادثين مستقلين إذا كان وقوع إحداهما أو عدم وقوعه لا يؤثر في وقوع الآخر. فمثلاً عند رمي قطعة عملة واحدة مرتين متتاليتين فإن نتيجة الرمية الثانية لا تتأثر بنتيجة الأولى.

جـ-الحوادث الشاملة (Exhaustive Events)

تسمي الحوادث A ، B ، C ... حوادث شاملة في تجربة ما إذا كان لابد من حدوث إحداها عند إجراء التجربة. فمثلاً عند اختيار طالب من الجامعة لمعرفة حالته ما إذا كان مدخنا أو غير مدخن تعتبر هذه الحالات حوادث شاملة لأنه لابد للفرد أن يكون له صفة واحدة من هذه الصفات. كذلك فإن الحصول على العدد 1 أو 2 أو 3 أو 4 أو 5 أو 6 عند رمي حجر النرد تعتبر حوادث شاملة لأنه لابد من حدوث إحداها.

مثال:

الجدول التالي يمثل توزيع عمال أحد المصانع حسب الحالة الاجتماعية للعامل والقسم الدي يعمل به:

الحالة الاجتماعية	أعزب	متزوج	المجموع
القسم الأول	5	7	12
القسم الثاني	8	14	22
القسم الثالث	10	6	16
المجموع	23	27	50

اختر عامل من الجدول السابق بطريقة عشوائية، ثم احسب الاحتمالات التالية:

- أن يكون أعزبا
- أن يكون متزوجا
- أن يكون من القسم الأول
- أن يكون من القسم الأول أو الثاني
- أن يكون من القسم الأول وأعزب

الحل:

نفرض أن الحادثة **A** أن يكون **العامل أعزب** أي A = {أن يكون العامل أعزب} فيكون الاحتمال المطلوب:

$$P(A)= \frac{\text{عدد العمال العزاب}}{\text{عدد العمال الكلي}} = 23/50= 0.46$$

نفرض أن الحادثة **B** أن يكون **العامل متزوج** أي أن B = {أن يكون العامل متزوج} فيكون الإحتمال المطلوب:

$$P(B)= \frac{\text{عدد العمال المتزوجين}}{\text{عدد العمال الكلي}} = 27/50= 0.54$$

نفرض أن الحادثة **C** أن يكون **العامل من القسم الأول** أي أن C = {أن يكون العامل من القسم الأول} فيكون الإحتمال المطلوب:

$$P(C)= \frac{\text{عدد عمال القسم الأول}}{\text{عدد العمال الكلي}} = 12/50= 0.24$$

تابع الحل:

نفرض أن الحادثة **D** أن يكون **العامل من القسم الأول أو الثاني** أي أن D = {أن يكون العامل من القسم الأول أو الثاني} فيكون الاحتمال المطلوب:

$$P(D)= \frac{\text{عدد عمال القسم الأول أو الثاني}}{\text{عدد العمال الكلي}} = (12+22)/50=34/50= 0.68$$

نفرض أن الحادثة **E** أن يكون **العامل من القسم الأول و أعزب** أي أن E = {أن يكون العامل من القسم الأول وأعزب} فيكون الاحتمال المطلوب:

$$P(E)= \frac{\text{عدد عمال القسم الأول وعزاب}}{\text{عدد العمال الكلي}} = 5/50= 0.1$$

مثال:-

الجدول التالي يمثل توزيع عمال أحد المصانع حسب الحالة الاجتماعية للعامل والقسم الذي يعمل به:

الحالة الاجتماعية	أعزب	متزوج	المجموع
القسم الأول	5	7	12
القسم الثاني	8	14	22
القسم الثالث	10	6	16
المجموع	23	27	50

اختير عامل من الجدول السابق بطريقة عشوائية، احسب الاحتمالات التالية:

- احتمال أن يكون العامل من القسم الأول أو الثاني.
- احتمال أن يكون العامل متزوجا أو من القسم الأول
- احتمال أن يكون العامل من القسم الثالث أو أعزب

الحل:

نفرض أن الحادثة **A1** أن يكون **العامل من القسم الأول** أي أن A1 = {أن يكون العامل من القسم الأول}
نفرض أن الحادثة **A2** أن يكون **العامل من القسم الثاني** أي أن A2 = {أن يكون العامل من القسم الثاني}
فيكون الاحتمال المطلوب:

$P(A1)=12/50$
$P(A2)=22/50$
$P(A1 \cup A2) = P(A1)+P(A2) = (12/50)+(22/50) = 34/50 = 0.68$

نفرض أن الحادثة **A1** أن يكون **العامل متزوجا** أي أن A1 = {أن يكون العامل متزوج}
نفرض أن الحادثة **A2** أن يكون **العامل من القسم الأول** أي أن A2 = {أن يكون العامل من القسم الأول}
فيكون الاحتمال المطلوب:

$P(A1)=27/50$
$P(A2)=12/50$
$P(A1 \cup A2) = P(A1)+P(A2)-P(A1 \cap A2) = (27/50)+(12/50)-(7/50) = 32/50 = 0.64$

تابع الحل:

نفرض أن الحادثة **A1** أن يكون **العامل من القسم الثالث** أي أن A1 = {أن يكون العامل من القسم الثالث}

نفرض أن الحادثة **A2** أن يكون **العامل أعزب** أي أن A2 = {أن يكون العامل أعزب}

فيكون الاحتمال المطلوب:

$P(A1)=16/50$

$P(A2)=23/50$

$$\mathbf{P(A1 \cup A2)} = P(A1)+P(A2)-P(A1 \cap A2) = (16/50)+(23/50)-(10/50) = 29/50 = 0.58$$

مثال:

الجدول التالي يمثل توزيع عمال أحد المصانع حسب الحالة الاجتماعية للعامل والقسم الدي يعمل به:

الحالة الاجتماعية	أعزب	متزوج	المجموع
القسم الأول	5	7	12
القسم الثاني	8	14	22
القسم الثالث	10	6	16
المجموع	23	27	50

اختير عامل من الجدول السابق بطريقة عشوائية، احسب الاحتمالات التالية:

- **احسب احتمال أن يكون العامل من القسم الأول بشرط أنه متزوج؟**
- **احتمال أن يكون العامل أعزب بشرط أنه من القسم الثالث؟**

الحل:

نفرض أن A1={أن يكون العامل من القسم الأول}

A2={أن يكون العامل متزوج}

B3={أن يكون العامل من القسم الثالث}

B4={أن يكون العامل أعزب}

فيكون بالتالي:

1- احتمال أن يكون العامل من القسم الأول بشرط أنه متزوج هو:

$$\frac{\text{احتمال أن يكون من القسم الأول ومتزوج}}{\text{احتمال أن يكون متزوج}}$$

$$P(A1 \mid A2) = \frac{P(A_1 \cap A_2)}{P(A_2)} = \frac{\frac{7}{50}}{\frac{27}{50}} = \frac{7}{27}$$

إذا احتمال أن يكون العامل من القسم الأول بشرط أنه متزوج هو 0.259

2- احتمال أن يكون العامل أعزب بشرط أنه من القسم الثالث هو:

$$\frac{\text{احتمال أن يكون العامل أعزب بشرط أنه من القسم الثالث}}{\text{احتمال أن يكون من القسم الثالث}}$$

$$P(B1 \mid B2) = \frac{P(B_1 \cap B_2)}{P(B_2)} = \frac{\frac{10}{50}}{\frac{16}{50}} = \frac{10}{16}$$

إذا احتمال أن يكون العامل أعزب بشرط أنه من القسم الثالث هو 0.625

مثال:

الجدول التالي يمثل توزيع عمال أحد المصانع حسب الحالة الاجتماعية للعامل والقسم الذي يعمل به فإذا سحب عاملان من المصنع مع الإرجاع (أي إرجاع العامل الأول قبل سحب العامل الثاني) احسب :

الحالة الاجتماعية	أعزب	متزوج	المجموع
القسم الأول	5	7	12
القسم الثاني	8	14	22
القسم الثالث	10	6	16
المجموع	23	27	50

اختير عاملان من الجدول السابق بطريقة عشوائية، احسب الاحتمالات التالية:

1. احتمال أن يكون العاملان من القسم الأول؟
2. احتمال أن يكون العاملان متزوجان؟
3. احتمال أن يكون للعاملين نفس الحالة الاجتماعية؟
4. احتمال أن يكون العاملان من القسم نفسه؟

الحل:

1- احتمال أن يكون العاملان من القسم الأول يعني أن يكون العامل الأول من القسم الأول (الحادثة A_1) والعامل الثاني من القسم الأول (الحادثة A_2) وحيث أنهما مستقلان (لأن السحب مع الإرجاع) فإن:

$$P(A_1 \cap A_2) = P(A_1) \times (P(A_2) = \frac{12}{50} \times \frac{12}{50} = \frac{144}{2500} = 0.0576$$

2- احتمال أن يكون العاملان متزوجان ، يعني أن يكون العامل الأول متزوج (الحادثة B_1) والعامل الثاني متزوج (الحادثة B_2) وحيث أنهما مستقلان (لأن السحب مع الإرجاع) فإن:

$$P(B_1 \cap B_2) = P(B_1) \times (P(B_2) = \frac{27}{50} \times \frac{27}{50} = \frac{729}{2500} = 0.2916$$

3- احتمال أن يكون للعاملين نفس الحالة الاجتماعية يعني أن يكون العاملان كلاهما متزوجين (الحادثة A) أو أن يكون كلاهما أعزبين (الحادثة B) فإن:

$$\begin{aligned} P(A \cup B) &= P(A) + (P(B) \\ &= P(A_1 A_2) + P(B_1 B_2) \\ &= P(A_1 \times A_2) + P(B_1 \times B_2) \\ &= \left[\frac{27}{50} \times \frac{27}{50}\right] + \left[\frac{23}{50} \times \frac{23}{50}\right] = 0.5032 \end{aligned}$$

4- احتمال أن يكون العاملان من القسم نفسه يعني أن يكون العاملان كلاهما من القسم الأول (الحادثة A) أو أن يكون كلاهما من القسم الثاني (الحادثة B) أو أن يكون كلاهما من القسم الثالث (الحادثة C) فإن:

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + (P(B) + (P(C) \\ &= P(A_1 A_2) + P(B_1 B_2) + P(C_1 C_2) \\ &= P(A_1 \times A_2) + P(B_1 \times B_2) + P(C_1 \times C_2) \\ &= \left[\frac{12}{50} \times \frac{12}{50}\right] + \left[\frac{22}{50} \times \frac{22}{50}\right] + \left[\frac{16}{50} \times \frac{16}{50}\right] = 0.3536 \end{aligned}$$

مثال:-

مصنع يقوم بإنتاج سلعة معينة به ثلاث آلات، تنتج الآلة الأولى 20% من إجمالي إنتاج السلعة وتنتج الآلة الثانية نسبة 35% والثالثة بنسبة 45% ، فإذا كانت نسبة الإنتاج المعيب في الثلاث آلات على الترتيب هو 2% و 2.5% و 3% ، سحبت وحدة عشوائيا من إنتاج المصنع فوجد أنها معيبة، احسب الاحتمالات التالية:

1- أن تكون القطعة المعيبة من إنتاج الآلة الأولى؟

2- أن تكون القطعة المعيبة من إنتاج الآلة الثانية؟

الحل:

نفرض أن

A1={إنتاج الآلة الأولى} $P(A1)=0.2$

A2={إنتاج الآلة الثانية} $P(A2)=0.35$

A3={إنتاج الآلة الثالثة} $P(A3)=0.45$

B= {إنتاج سلعة معينة}

فيكون بالتالي:

$P(B|A1)=0.02$

$P(B|A2)=0.025$

$P(B|A3)=0.03$

تابع الحل:-

تكون السلعة من إنتاج الآلة الأولى إذا علم - بشرط - أنها معيبة هو:

$$P(A_1|B)=\frac{P(A_1)P(B|A_1)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)}=\frac{0.2\times 0.02}{(0.2\times 0.02)+(0.35\times 0.025)+(0.45\times 0.03)}=0.152$$

واحتمال أن تكون السلعة من إنتاج الآلة الثانية إذا علم - بشرط - أنها معيبة هو:

$$P(A_2|B)=\frac{P(A_2)P(B|A_2)}{\sum_{i=1}^{n}P(A_i)P(B|A_i)}=\frac{0.35\times 0.025}{(0.2\times 0.02)+(0.35\times 0.025)+(0.45\times 0.03)}=0.333$$

المتغير العشوائي Random Variable:

المتغير العشوائي هو الذي يأخذ قيما حقيقية مختلفة تعبر عن نتائج فراغ العينة، ومن ثم مجال هذا المتغير، يشمل كل القيم الممكنة له، ويكون لكل قيمة من القيم التي يأخذها المتغير احتمال معين .

وينقسم المتغير العشوائي إلى قسمين هما:

- المتغيرات العشوائية المنفصلة Discrete Random Variables
- المتغيرات العشوائية المتصلة(المستمرة) Continuous Random Variables

مثال :-

إذا كان إحتمال نجاح الطالب في مقرر التحليل الإحصائي 80% تم إختيار 4 طلاب المطلوب :-

1. كون جدول توزيع ثنائي الحدين .
2. أوجد إحتمال نجاح 3 طلاب .
3. أوجد إحتمال رسوب 3 طلاب .
4. أوجد إحتمال نجاح طالبين على الأقل .
5. القيمة المتوقعة (الوسط الحسابي) .
6. الانحراف المعياري .

الحل :-

P = 0.80 , (1-P= 0.20) , n=4

1- جدول توزيع ثنائى الحدين :-

عدد الطلاب الناجحين	عدد الطلاب الراسبين	الاحتمال	الناتج
0	4	$= 4C0 \times (0.80)^0 \times (0.20)^4$	0.0016
1	3	$= 4C1 \times (0.80)^1 \times (0.20)^3$	0.0256
2	2	$= 4C2 \times (0.80)^2 \times (0.20)^2$	0.1536
3	1	$= 4C3 \times (0.80)^3 \times (0.20)^1$	0.4096
4	0	$= 4C4 \times (0.80)^4 \times (0.20)^0$	0.4096

تابع الحل :-

2- أوجد إحتمال نجاح 3 طلاب :-

P(3) = 0.4096

3- أوجد إحتمال رسوب 3 طلاب :-

P(1) = 0.0256

4- أوجد إحتمال نجاح طالبين على الأقل :-

P(2)+P(3)+P(4) = 0.9728

5- القيمة المتوقعة (الوسط الحسابي) =

$\mu = n \times p = 4 \times 0.80 = 3.2$

6- الانحراف المعياري =

$$\sigma = \sqrt{n \times p \times (1-p)} = \sqrt{4 \times 0.8 \times 0.2} = 0.8$$

تابع توزيع ثنائي الحدين :-

مثال:-

وجد في إنتاج أحد المصانع أنه من بين 1000 وحدة إنتاج يوجد 150 وحدة معيبة. أخذت عينة بإرجاع مكونة من 5 وحدات، أوجد الاحتمالات التالية:

1– الوحدات المختارة كلها سليمة

2– على الأكثر توجد واحدة معيبة

3– على الأقل توجد وحدتان معيبتان

4– القيمة المتوقعة و التباين للوحدات المعيبة .

الحل

احتمال النجاح (الحصول على وحدة معيبة) $p = 150/1000 = 0.15$

احتمال الفشل (عدم الحصول على وحدة معيبة) $q = 1-p= 1-0.15 =0.85$

عدد المحاولات (عينة بإرجاع مكونة من 5 وحدات) $n = 5$

X متغير عشوائي يمثل عدد الوحدات المعيبة يأخذ القيم $0, 1, 2, 3, 4, 5$

ويكون له توزيع ذي الحدين:

تابع توزيع ثنائي الحدين :-

$$p(X=x)=\binom{5}{x}(0.15)^{x}(0.85)^{5-x} \quad , \quad x=0,1,2,3,4,5$$

1- الوحدات كلها سليمة يعنى أن X = 0

$$p(X=0)=\binom{5}{0}(0.15)^{0}(0.85)^{5-0} = \frac{5!}{0!(5-0)!}(1)(0.85)^{5} = 0.4437$$

2- على الأكثر توجد وحدة معيبة يعنى أن $X \le 1$

$$P(X\le1) = p(X=0) + p(X=1)$$

$$p(X\le1)=\binom{5}{0}(0.15)^{0}(0.85)^{5}+\binom{5}{1}(0.15)^{1}(0.85)^{4}$$

$$= 0.4437 + \frac{5!}{1!5!}(0.15)(0.522)$$

$$= 0.4437 + 5\times0.0783 = 0.4437 + 0.3915 = 0.8352$$

تابع توزيع ثنائي الحدين :-

$$p(X=x)=\binom{5}{x}(0.15)^x(0.85)^{5-x} \quad , \quad x=0,1,2,3,4,5$$

3- على الأقل توجد وحدتان معيبتان يعن أن $X \geq 2$

$$P(X \geq 2) = 1 - p(X < 2)$$
$$= 1 - [p(X = 0) + p(X = 1)]$$
$$= 1 - 0.8325 = 0.1648$$

4 – القيمة المتوقعة و التباين للوحدات المعيبة .

القيمة المتوقعة $= n \,.\, P = 5 \times 0.15 = 0.75$

التباين $= n \times p \times (1-p) =$

$= 5 \times 0.15 \times 0.85 = 0.6375$

مثال:-

إذا كان من المعلوم أن عدد الوحدات التي تستهلكها الأسرة من سلعة معينة خلال الشهر تتبع توزيع بواسون بمتوسط 3 وحدات شهريا، إذا عرف المتغير العشوائي x بأنه عدد الوحدات التي تستهلكها الأسرة خلال الشهر من هذه السلعة.

المطلوب:

- ما نوع المتغير العشوائي؟
- اكتب شكل دالة الاحتمال لهذا المتغير.
- احسب الاحتمالات التالية:
  - احتمال أن الأسرة تستهلك وحدتين خلال الشهر؟
  - احتمال أن أسرة ما تستهلك 3 وحدات على الأكثر خلال الشهر؟
- احسب الوسط الحسابي، والانحراف المعياري لعدد الوحدات المستهلكة.
- حدد شكل التوزيع.

الحل:-

عدد الوحدات التي تستهلكها الأسرة متغير كمي منفصل، ومدى هذا المتغير في هذه الحالة هو:

$$X : \{x = 0,1,2,3,...\}$$

شكل دالة الاحتمال:

بما أن متوسط عدد الوحدات التي تستهلكها الأسرة خلال الشهر هو: $\mu = 3$ ، إذا دالة الاحتمال هي:

$$P(x) = \frac{e^{-\mu}\mu^{x}}{x!}$$

$$= \frac{e^{-3}3^{x}}{x!} \quad , \quad x = 0,1,2,...$$

حساب الاحتمالات:

حساب احتمال أن أسرة ما تستهلك وحدتين خلال الشهر، p(2)

$$P(2) = \frac{e^{-3}3^{2}}{2!} = \frac{0.0498 \quad (9)}{2 \times 1} = 0.22404$$

تابع الحل:-

احتمال أن أسرة ما تستهلك 3 وحدات على الأكثر خلال الشهر هو:

$$P(X \leq 3) = p(3) + p(2) + p(1) + p(0)$$

$$= \left[\frac{3^{3}}{3!} + \frac{3^{2}}{2!} + \frac{3^{1}}{1!} + \frac{3^{0}}{0!}\right]\left[\frac{0.0498}{1}\right]$$

$$= [0.0498]\left(\frac{27}{6} + \frac{9}{2} + \frac{3}{1} + \frac{1}{1}\right) = 0.0498 \quad (13) = 0.6474$$

حساب الوسط الحسابي، والانحراف المعياري لعدد حالات الاستجابة:

الوسط الحسابي (μ) في حالة التوزيع بواسون هو معلمة معطاة هي: $\mu = 3$

في هذا التوزيع، فإن التباين يساوي الوسط الحسابي: أي أن: $\sigma^2 = \mu = 3$

ومن ثم يكون الانحراف المعياري هو: $\sigma = \sqrt{\mu} = \sqrt{3} = 1.732$

ويمكن حساب معامل الاختلاف النسبي، بتطبيق المعادلة التي سبق استخدامها ، وهو:

$$C.V = \frac{\sigma}{\mu} \times 100 = \frac{1.732}{3} \times 100 = 57.7\%$$

تحديد شكل التوزيع:

<u>دائما توزيع بواسون موجب الالتواء</u>

مثال:-

يتلقى قسم شرطة في المتوسط **5** مكالمات في الساعة فيكون احتمال تلقى مكالمتين في ساعة مختارة عشوائياً هو :

$$P(x)=\frac{e^{-\mu}\mu^{x}}{x!}$$

$$=\frac{e^{-5}5^{x}}{x!}= \ , \quad x=0,1,2,\ldots$$

$$=\frac{(25)(0.00674)}{(2)(1)}=0.08425$$

التوزيع الطبيعي القياسي (المعيارى) :-

- احتمال وقوع أية مشاهدة على بعد انحراف معياري واحد من الوسط الحسابي هو 0.6827
- احتمال وقوع أي مفردة على بعد إنحرافين معياريين من الوسط الحسابي هو 0.9545
- احتمال وقوع أية مفردة على بعد ثلاثة انحرافات معيارية من الوسط الحسابي هو 0.9973

والشكل التالي يوضح ذلك:

العلاقة بين التوزيع الطبيعي والتوزيع الطبيعي القياسي

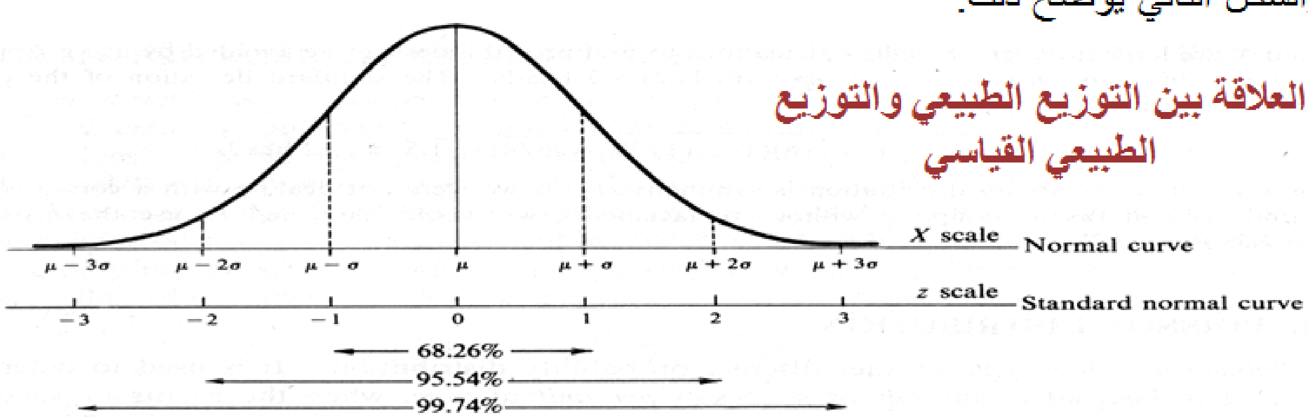


Fig. 3-4

مثال :-

تم دراسة متوسط طول الطالب في كلية إدارة الأعمال هو 180 سم و ذلك بانحراف معياري 10 سم تم اختيار أحد الطالب عشوائياً فإذا علمت أن هذه الظاهرة تتبع التوزيع الطبيعي فأوجد :-

1-احتمال أن ينحصر طول الطالب بين 170 سم و 190 سم ((p(170<x<190).

2-احتمال أن ينحصر طول الطالب بين 160 سم و 200 سم ((p(160<x<200).

3-احتمال أن ينحصر طول الطالب بين 150 سم و 210 سم ((p(150<x<210).

4- احتمال أن يكون طول الطالب أقل من 190 سم ((p(x<190).

5- احتمال أن يكون طول الطالب أكبر من 190 سم ((p(x>190).

6-احتمال أن يكون طول الطالب أكبر من 150 سم ((p(x>150).

7-احتمال أن يكون طول الطالب أقل من 160 سم ((p(x<160).

الحل :-

1-احتمال أن ينحصر طول الطالب بين 170 سم و 190 سم ((p(170<x<190):-

$$z = \frac{x-\mu}{\sigma}$$

$$\frac{170-180}{10} < z < \frac{190-180}{10}$$

$$-1 < z < 1$$

$$P = 68.26\%$$

الحل :-

2-احتمال أن ينحصر طول الطالب بين 160 سم و 200 سم -: (p(160<x<200))

$$z = \frac{x - \mu}{\sigma}$$

$$\frac{160 - 180}{10} < z < \frac{200 - 180}{10}$$

$$-2 < z < 2$$

P =95.45%

الحل :-

3-احتمال أن ينحصر طول الطالب بين 150 سم و 210 سم -:(p(150<x<210))

$$z = \frac{x - \mu}{\sigma}$$

$$\frac{150 - 180}{10} < z < \frac{210 - 180}{10}$$

$$-3 < z < 3$$

P =99.74%

الحل :-

4- احتمال أن يكون طول الطالب أقل من 190 سم (p(x<190)) :-

$$z = \frac{x - \mu}{\sigma}$$

$$z < \frac{190 - 180}{10}$$

$$z < 1$$

P = (0.6826/2) + 0.5 = 84.13%

الحل :-

5- احتمال أن يكون طول الطالب أكبر من 190 سم (p(x>190)):-

$$z = \frac{x - \mu}{\sigma}$$

$$z > \frac{190 - 180}{10}$$

$$z > 1$$

P = 0.5- (0.6826/2) = 15.87%

الحل :-

6-احتمال أن يكون طول الطالب أكبر من 150 سم ((p(x>150)) :-

$$z = \frac{x - \mu}{\sigma}$$

$$z > \frac{150 - 180}{10}$$

z > -3

z < 3

P = (0.9974/2)+0.5 = 99.87%

الحل :-

7- احتمال أن يكون طول الطالب أقل من 160 سم ((p(x<160)):-

$$z = \frac{x - \mu}{\sigma}$$

$$z < \frac{160 - 180}{10}$$

z > -2

z > 2

P = 0.5- (0.9545/2) = 0.02275= 2.275 %

الاستدلال الإحصائى :-

تهتم نظرية العينات بدراسة العلاقة بين المجتمع والعينات المسحوبة منه فيما يسمى **بالاستدلال الإحصائي statistical inference**

يعتبر **الاستدلال الإحصائي** من أهم الأدوات المساعدة على اتخاذ القرارات في الاقتصاد والأعمال والعلوم، ويشمل الاستدلال الإحصائي اختبار الفرضيات والتقدير.

ولكى يكون التقدير (واختبار الفروض) سليمً، ينبغي أن يبنى على عينة ممثلة للمجتمع، ويمكن تحقيق ذلك **بالمعاينة العشوائية**، حيث يكون لكل مفردة في المجتمع فرصة متكافئة للدخول في العينة .

العينة العشوائية:-

وهناك عدة طرق لأخذ العينات من المجتمع لاستخدامها في الاستدلال الإحصائي ومن أشهر هذه الطرق هي **العينة العشوائية** وهي العينة التي تكون لكل مفرده من مفردات المجتمع نفس فرصة الاختيار في العينة.

فمثلاً نستعين بعينه مسحوبة من المجتمع لتقدير معالم هذا المجتمع مثل متوسطة أو تباينه أو غير ذلك. أو أعطاء عينه من المرضى بارتفاع الضغط، مثلاً دواء معين ثم قياس ضغطهم قبل وبعد تناولهم لهذا الدواء لمعرفة ما إذا كان هذا الدواء مفيد في خفض الضغط أم لا.

المجتمع Population

أي مجموعات من المفردات تشترك في صفه أو صفات وتكون موضوع دراسة أو بحث فإن هذه المجموعة يطلق عليها إحصائيا مجتمع الدراسة أو اختصاراً **المجتمع Population**.

والمجتمع قد يكون مجموعة ما من البشر أو أشجار أنواع معينه من الفاكهة أو الحيوانات الزراعية أو إنتاج دولة ما لسلع معينه خلال فترة زمنية محدده...الخ.

والمجتمع قد يكون محدوداً إذا كان يمكن حصر عدد أفراده مثل سكان مدينة ما أو طلاب مرحلة تعليمية معينة

وقد يكون المجتمع غير محدود (لانهائي) إذا كان لا يمكن حصر عدد أفراده مثل النجوم والكواكب أو الكائنات الحية بمياه المحيطات والأنهار

وعند دراسة صفة ما أو صفات معينه لمجتمع ما فإن البيانات الإحصائية عن تلك الصفة أو الصفات تجمع بأحد أسلوبين:-

أولاً: أسلوب الحصر الشامل (census): وفيه تجمع البيانات عن كل مفرده من مفردات المجتمع، وهذا الأسلوب لا يتبع عادة إلا في حالة التعدادات التي تجريها الدول وتدعمها بإمكانيات ضخمه مثل تعدادات السكان والتعدادات الصناعية والتعدادات الزراعية.

الثاني: أسلوب المعاينة (Sampling method): وفيه يتم جمع البيانات عن جزء من مفردات المجتمع يختار بطريقة أو بأخرى ويطلق عليه عينه (Sample) ثم بعد ذلك يتم تعميم نتائج الدراسة على المجتمع بأكمله.

أقسام العينات:-

تنقسم العينات عادة إلى قسمين رئيسين وهما عينات عشوائية وعينات غير عشوائية، وفيما يلي تفصيل لكل قسم منها:

1. العينات العشوائية:

وهي تلك العينات التي يتم اختيار مفرداتها حسب خطه إحصائية لا يكون فيها للباحث أو لمفردات العينة دخل في اختيار أي مفرده فيها ، حيث يتم الاختيار باستخدام أساليب معينة تلعب الصدفة خلالها الدور الأول في اختيار المفردة ولكن بشرط أن يتحقق لجميع المفردات احتمال ثابت ومحدد للاختيار. والعينات العشوائية إذا ما تم اختيارها بالطريقة العلمية السليمة والمناسبة يمكن أن تكفل درجه عالية من دقة التمثيل للمجتمعات المسحوبة منها لذلك فهي الوسيلة الأساسية في حالة البحوث العلمية الدقيقة

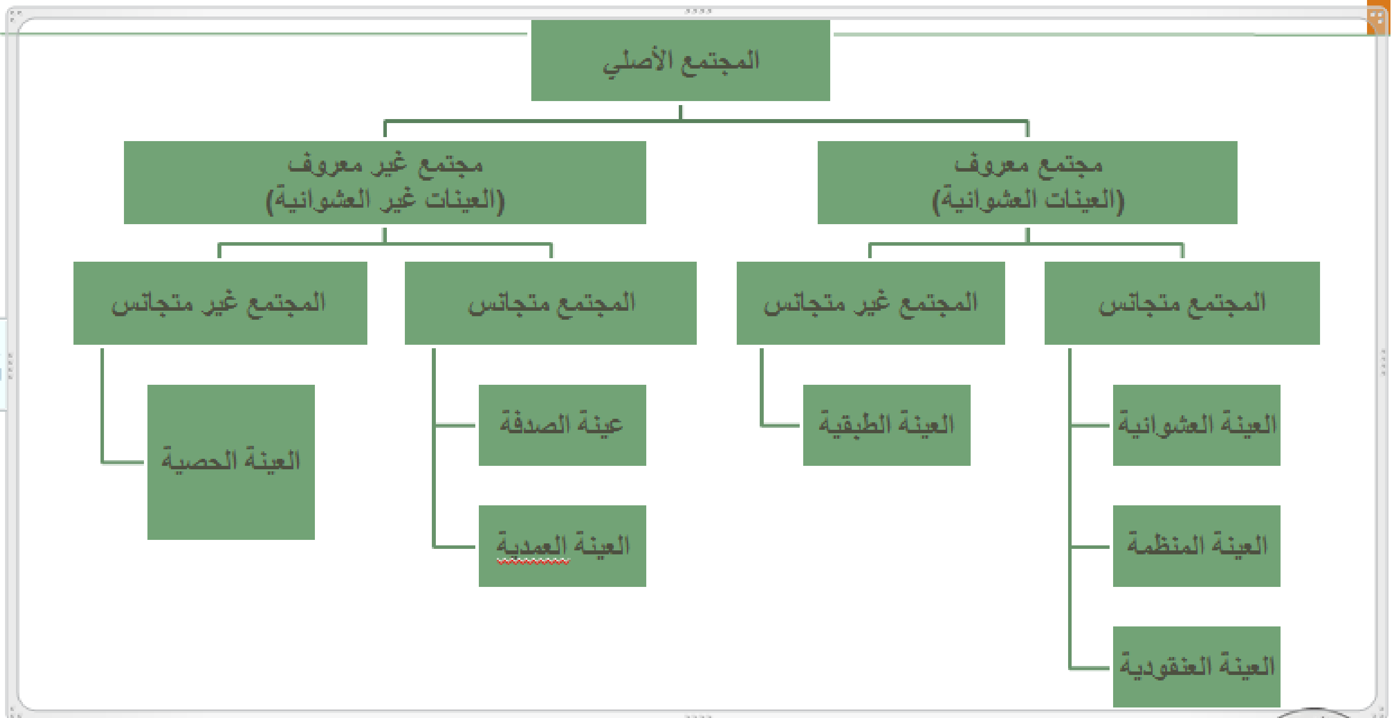
تابع أقسام العينات:-

2. العينات غير العشوائية:

وهي تلك العينات التي لا تكفل لجميع مفردات المجتمع احتمال ثابت ومحدد للاختيار، وغالباً يتدخل الباحث في عملية الاختيار بصورة أو بأخرى.

وسيتم فيما يلي استعراض لأهم أنواع العينات العشوائية والعينات غير العشوائية.

أقسام العينات:



أ ـ العينات الاحتمالية:

العينة العشوائية	جميع عناصر المجتمع لها نفس الفرصة في الظهور في العينة
العينة الطبقية	يقسم المجتمع إلى طبقتين على الأقل ثم نختار العينة من كل منهما
العينة المنتظمة	نختار نقطة بداية من المجتمع ثم نختار العنصر الموجود على بعد ثابت من هذه النقطة
العينة العنقودية	يقسم المجتمع إلى مساحات أو أجزاء ثم نختار عشوائيا بعض هذه المساحات، ثم نختار جميع عناصرها بالعينة.

ب - العينات غير الاحتمالية:

عينة الصدفة	يتم اختيارها عن طريق الصدفة
العينة العمدية (القصدية)	يتم اختيار أفراد العينة تحت شروط معينة لتحقيق الهدف من التجربة
العينة الحصية	يقسم المجتمع إلى أجزاء ثم نختار العينة من كل جزء من أجزاء المجتمع وفقا للنسب المحددة

أخطاء البيانات الإحصائية: -

تتعرض البيانات الإحصائية التي يتم جمعها إلى نوعين من الأخطاء:

1. **خطأ التميز أو التحيز:** وهو ذلك الخطأ الناتج عن مصادر متعددة، منها أخطاء في تصميم البحث أو التجربة أو أخطاء فنيه أثناء جمع البيانات أو خلال العمليات الحسابية التي تتم على البيانات المتجمعة.. أخطاء التميز تزداد بازدياد الفروق بين الإمكانيات

(المادية والفنية) اللازم توافرها لضمان أقصى درجة دقة ممكنه وبين الإمكانيات الفعلية المتاحة للباحث.

1. **خطأ المعاينة العشوائية أو خطأ الصدفة:** وهو الخطأ الناتج عن فروق الصدفة بين مفردات المجتمع التي دخلت العينة وبين تلك المفردات التي لم تشأ الصدفة أن تدخل العينة

وفيما يلي شرح لهذين الخطأين:

الخطأ المعياري:

بمعرفة قيمة الانحراف المعياري لقيم العينة يمكن تقدير قيمة الخطأ المعياري في الانحراف المعياري للعينة باعتماد المعادلة الآتية:-

$$SE = \left[\frac{S}{\sqrt{n}} \right] \left[\sqrt{1 - (\frac{n}{N})} \right]$$

حيث ان:

(N) = حجم مجتمع العينة و (n) = حجم العينة

تابع الخطأ المعياري:

وفق هذه المعادلة تؤخذ نسبة العينة إلى مجتمعها، وكلما كبرت هذه النسبة تحسن تمثيلها لمجتمع الدراسة، أما عندما يكون حجم مجتمع العينة مجهولا، حينها تعتمد المعادلة الآتية:-

$$SE = \frac{S}{\sqrt{n}}$$

الخطأ المعياري = الانحراف المعياري لقيم العينة / جذر حجم العينة

تابع الخطأ المعياري:

أما عندما يكون حجم العينة أكثر من (100) فتعتمد المعادلة أدناه:-

$$SE = \frac{S}{\sqrt{n^2} - 1}$$

الخطأ المعياري = الانحراف المعياري لقيم العينة / جذر (حجم العينة)^2-1

تابع مستوى الثقة وحدودها:

- مستوى ثقة إحصائية قدره (68.26%) أو باحتمالية قدرها (0.6827) يقع متوسط مجتمع الدراسة بين قيمة متوسط متوسطات العينات و (+ و -) درجة واحدة من الخطأ المعياري .
- مستوى ثقة إحصائية قدرها (95%)، أو باحتمالية (0.95) يقع متوسط مجتمع الدراسة بين متوسط متوسطات العينات و (+ و –) درجتان من الخطأ المعياري تقريبا.

تابع مستوى الثقة وحدودها:

مستوى ثقة إحصائية قدرها (99%) أو باحتمالية قدرها (0.99) يقع متوسط مجتمع الدراسة بين قيمة متوسط متوسطات العينات و (+ و –) ثلاث درجات من قيمة الخطأ المعياري تقريبا.

وتسمى هذه بمستويات الثقة Confidence Level و يعبر عنها بإشارة النسبة المئوية (%) بان تكون التقديرات صحيحة أو باحتمالية (0.01) أو (0.05) أن تكون خاطئة.

تابع حجم العينة:-

(ب) توزيع (ت) للقيم :

من الضروري اخذ الحذر عندما يكون حجم العينة صغيرا اقل من (30) وذلك لأنها تتطلب إجراءات خاصة عند التحليل. فعندما يكون حجم العينة أكثر من (30) يتجه توزيع قيمها نحو التوزيع الطبيعي وبغض النظر عن التوزيع الحقيقي لقيم مجتمع الدراسة.

وبالنسبة للعينات الصغيرة الحجم فان توزيع قيمها يتأثر بطبيعة توزيع قيم المجتمع المأخوذة منه. **وعندما يكون توزيع قيم المجتمع معروفا أو متوقعا أن لا يكون طبيعي حينها يجب اعتماد حجم كبير للعينة**. أما إذا كان توزيع قيم المجتمع طبيعيا عندها يمكن اخذ عينات بحجم صغير ويعتمد توزيع (ت) (T) في التحليل و المقارنة.

العوامل المحددة لحجم العينة:

درجة التباين في خصائص مجتمع الدراسة: يلعب التباين في خصائص مجتمع الدراسة دورا مهما في تحديد درجة دقة نتائج العينة، فكلما كان التباين كبيرا تطلب الأمر زيادة حجم العينة ليكون تمثيلها للتباين في المجتمع صحيحا.

طريقة التحليل المعتمدة: عند إقرار حجم العينة، من الضروري تحديد الحجم الأصغر المقبول للعينة في المجاميع الثانوية ضمن مجتمع الدراسة ، إذ أن بعض الاختبارات الإحصائية تتطلب عددا معينا كحد أدنى لكل فئة أو صنف لتكون النتائج ذات معنى.

تقدير الوسط الحسابي للمجتمع : -

أ – احسب الوسط الحسابي للعينة $\bar{x}$.

ب- احسب الخطأ المعياري للوسط والذي يساوي: $\frac{\sigma}{\sqrt{n}}$

جـ - أضرب الخطأ المعياري للوسط في معامل الثقة المناسب (أو الدرجة المعيارية) حسب درجة الثقة المطلوبة أي أحسب: $Z \frac{\sigma}{\sqrt{n}}$

د- فعندما نطرح حاصل الضرب السابق من الوسط الحسابي للعينة نحصل بالتالي على الحد الأدنى لفترة التقدير، وعندما نجمع حاصل الضرب مرة أخرى على الوسط الحسابي للعينة نحصل بالتالي على الحد الأعلى لفترة التقدير.

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

تابع تقدير الوسط الحسابي للمجتمع : -

درجة الثقة	معامل الثقة Z
68.26%	1
90%	**1.65**
95 %	**1.96**
95.44%	2
99%	**2.58**

مع ملاحظة أنه إذا كان الانحراف المعياري للمجتمع غير معروف - وهو غالبا ما يحدث في الواقع – فيمكن استخدام الانحراف المعياري للعينة S بدلا منه طالما كان حجم العينة كبيرا بدرجة كافية وتصبح فترة تقدير الوسط الحسابي للمجتمع كما يلي :

$$\hat{\mu} = \overline{X} \pm Z \frac{S}{\sqrt{n}}$$

ولإيضاح هذه النقطة بشيء من التفصيل نأخذ المثال التالي :

مثال : -

لو أردنا معرفة متوسط الدخل اليومي لمجموعة من الناخبين في دولة ما، فإن ذلك يبدو أمرا صعباً من الناحية العملية نظراً لكبر حجم مجتمع الناخبين، إضافة إلى طول الوقت والتكاليف. لذا فإن الأسلوب العلمي المتبع في حالة كهذه هو اختيار عينة عشوائية نستطيع من خلال معرفة نتائجها لتقدير متوسط دخول الناخبين في هذه الدولة.

فلو سحبت عينة عشوائية من مجموع مجتمع الناخبين في دولة ما حجمها 100 ناخب، فإذا كان الوسط الحسابي والانحراف المعياري للدخل السنوي للناخبين بالعينة هما على الترتيب 90 ألف ريال و 25 ألف ريال.

المطلوب:

أوجد فترة تقدير للوسط الحسابي للدخل السنوي لمجموع الناخبين في هذه الدولة بدرجة ثقة 95% ؟

الحل : -

بما أن فترة تقدير الوسط الحسابي للمجتمع هي : $\hat{\mu} = \bar{X} \pm Z \frac{S}{\sqrt{n}}$

والمعلومات المعطاة هي :

حجم العينة n = 100

الوسط الحسابي للعينة = $\bar{X} = 90$

والانحراف المعياري للعينة S = 25

وحيث أن درجة الثقة هي 95% فإن: Z = 1.96 حسب ما هو موضح في الجدول السابق. وبالتالي فإن فترة تقدير الوسط الحسابي للدخل السنوي لمجتمع الناخبين بدرجة ثقة 95% هي :

تابع الحل : -

$$\hat{\mu} = 90 \pm 1.96 \frac{25}{\sqrt{100}}$$
$$= 90 \pm 1.96 (2.5)$$
$$= 90 \pm 4.9$$

$$\mu = \begin{cases} 85.1 \\ 94.9 \end{cases}$$

أي أن الوسط الحسابي للدخل السنوي لمجتمع الناخبين يتراوح بين 85.1 ألف ريال كحد أدنى، 94.9 ألف ريال كحد أعلى، وذلك بدرجة ثقة 95 %.

تحديد حجم العينة لتقدير الوسط الحسابي للمجتمع :-

يعتبر تحديد حجم العينة المناسب من المشاكل المهمة والشائعة التي تواجه الباحثين في مختلف المجالات، وبالذات عند دراسة الظواهر السياسية والإجتماعية ...الخ ، ويختلف تحديد حجم العينة باختلاف الهدف من التقدير.
فإذا كان المطلوب هو تقدير الوسط الحسابي للمجتمع، فإن فترة تقدير الوسط هي كما سبق وأن أوضحنا :

$$\hat{\mu} = \bar{X} \pm Z \frac{\sigma}{\sqrt{n}}$$

ومنها نجد أن حجم العينة يأخذ الشكل التالي :

$$\boxed{n = \frac{Z^2 \sigma^2}{e^2}}$$

مثال : -

إذا كانت دخول الأفراد اليومية في إحدى دول العالم النامية تتبع التوزيع الطبيعي بانحراف معياري $\sigma = 15$ دولاراً، فما هو حجم العينة المناسب لتقدير متوسط دخول الأفراد في هذه الدولة بحيث لا يتعدى الخطأ في تقدير متوسط الدخل اليومي 5 دولارات، وذلك بدرجة ثقة 99 %؟

الحل : -

في هذا المثال نجد أن :

درجة الثقة 99 % أي أن : Z = 2.58

أقصى خطأ مسموح به هو 5 دولارات، أي أن : e = 5

والانحراف المعياري للمجتمع : $\sigma = 15$

وبالتعويض بهذه القيم في المعادلة التي تحدد حجم العينة وهي :

$$n = \frac{Z^2 \sigma^2}{(e)^2}$$

تابع الحل : -

فإن حجم العينة مقرباً لأقرب عدد صحيح هو :

$$n = \frac{(2.58)^2 (15)^2}{5^2}$$
$$= \frac{(6.65)(225)}{25}$$
$$= \frac{1496.25}{25} = 59.85 \approx 60$$

أي أنه يجب على الباحث أن يأخذ عينة لا يقل حجمها عن 60 فرداً حتى يكون لديه تقديراً دقيقاً عن متوسط دخول الأفراد في هذه الدولة بحيث لا يتعدى الخطأ في تقديره لمتوسط الدخل عن خمس دولارات، وذلك بدرجة ثقة 99 %.

مثال : -

يرغب أحد مدراء إحدى المصانع في تقدير متوسط عدد الدقائق التي يأخذها العمال لإنجاز عملية صناعية معينة بحيث لا يتعدى الخطأ في تقدير متوسط الأداء في حدود ± 3 دقيقة وبدرجة ثقة 90% . ويعلم المدير من خبرته الماضية أن الانحراف المعياري σ هو 15 دقيقة .

الحل : -

في هذا المثال نجد أن :

درجة الثقة 90% أي أن : Z = 1.65

أقصى خطأ مسموح به هو 3 دقائق، أي أن : e = 3

والانحراف المعياري للمجتمع : $\sigma = 15$

وبالتعويض بهذه القيم في المعادلة التي تحدد حجم العينة وهي :

$$n = \frac{Z^2 . \sigma^2}{(e)^2}$$

تابع الحل : -

فإن حجم العينة مقرباً لأقرب عدد صحيح هو :

$$n = \frac{(1.65)^2 (15)^2}{3^2}$$

$$= \frac{(2.72)(225)}{9}$$

$$= \frac{612}{9} = 68$$

أي أنه يجب على المدير أن يأخذ عينة لا يقل حجمها عن 68 فرداً حتى يكون لديه تقديراً دقيقاً لعدد الدقائق التى يأخذها العمال لإنجاز عملية صناعية معينة بحيث لا يتعدى الخطأ في تقديره لمتوسط الإنجاز عن ثلاث دقائق، وذلك بدرجة ثقة 90 %.

مثال : -

العينة أقل من 30 و المجتمع يتبع التوزيع الطبيعي :

سحبت عينة عشوائية من **n=10** بطارية فلاش متوسطها 5 ساعات، والانحراف المعياري للعينة **s=1** ساعة من خط إنتاج من المعروف أنه ينتج بطاريات عمرها موزع طبقاً للتوزيع الطبيعي .

المطلوب :

إيجاد فترة الـ **95%** ثقة للمتوسط غير المعلوم لعمر البطاريات في المجتمع كله.

تابع الحل :

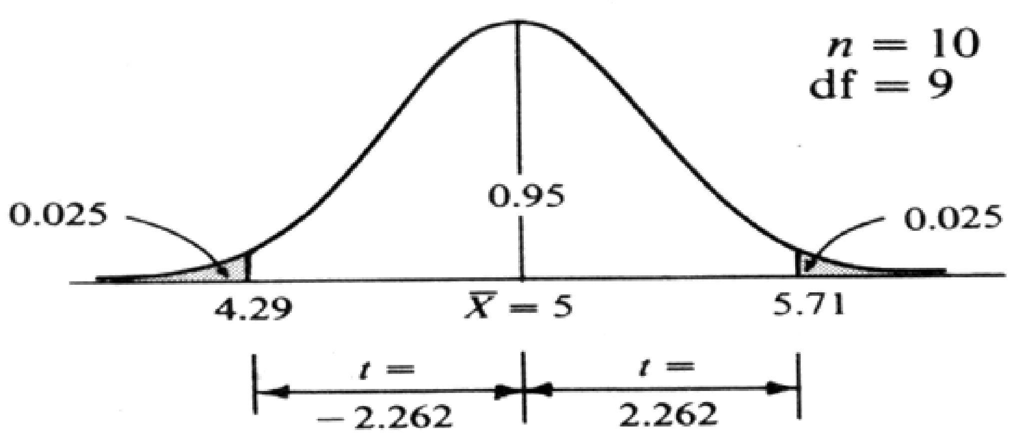
الحل:

لإيجاد فترة الـ **95%** ثقة للمتوسط غير المعلوم لعمر البطاريات فى المجتمع كله، فإننا نوجد أولا قيمة **t 0.025** و التى تكون معها **2.5%** من المساحة عند الأطراف لدرجات حرية **n-1=9** . ونحصل على هذه القيمة من خلال الرجوع إلى جدول **t** بالتحرك تحت عمود **0.025** حتى درجات حرية **9** والقيمة التى سيتم التحصل عليها هى **2.262** إذن:

$$\hat{\mu} = \overline{X} \pm 2.262 \frac{s}{\sqrt{n}} = 5 \pm 2.262 \frac{1}{\sqrt{10}} \cong 5 \pm 2.262(0.316) \cong 5 \pm 0.71$$

تابع الحل :

وتقع $\hat{\mu}$ بين **4.29 , 5.71** ساعة بدرجة ثقة **95%** (أنظر الشكل التالي):



مثال : -

عينة عشوائية حجمها 144 ناخباً سحبت من إحدى المدن فوجد أن عدد المؤيدين في العينة لمرشح معين هو 60 ناخباً، أنشئ فترة تقدير لنسبة المؤيدين لهذا المرشح في المدينة كلها بدرجة ثقة 95 %.

الحل:

نحسب أولاً نسبة المؤيدين للمرشح في العينة $\hat{P}$ التي نحصل عليها بقسمة عدد المؤيدين له على العدد الكلي للعينة (حجم العينة) أي أن :

$$\hat{P} = \frac{60}{144} = 0.42$$

تابع الحل : -

وحيث أن درجة الثقة المطلوبة هي 95 % فإن معامل الثقة المناسب هو: Z = 1.96 وفترة تقدير نسبة المؤيدين لهذا المرشح في المدينة تأخذ الشكل التالي :

$$P = \hat{P} \pm z\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$$

وبالتعويض عن حجم العينة n = 144

والنسبة في العينة $\hat{P} = 0.42$ ، $1-\hat{P} = 1-0.42 = 0.58$

ومعامل الثقة Z = 1.96

تابع الحل : -

نحصل بعدها على :

$$P = 0.42 \pm 1.96\sqrt{\frac{0.42 \times 0.58}{144}}$$

$$= 0.42 \pm (1.96)(0.0411)$$

$$= 0.42 \pm 0.08$$

$$\therefore \quad P \begin{cases} 0.34 \\ 0.50 \end{cases}$$

تابع الحل : -

أي أن نسبة المؤيدين للمرشح في المدينة تتراوح بين 0.34 , 0.50 وذلك بدرجة ثقة 95 % ، بمعنى آخر أن نسبة مؤيدي هذا المرشح في هذه المدينة لا تتجاوز 50 % كحد أعلى، وبالتالي ففرصته في الفوز كمرشح قد لا تكون كبيرة وذلك بدرجة ثقة 95 % بمعنى أن هذا الحكم لا تتجاوز نسبة الخطأ فيه 5 %.

الخطأ في اتخاذ القرار : -

الخطأ من النوع الأول : Type I error

الخطأ من النوع الأول هو "رفض الفرض العدمي بينما هو صحيح". أي أنه على الرغم من أن الفرض العدمي في الواقع صحيح وكان من الواجب قبوله فقد تم أخذ قرار خاطئ برفضه. وباختصار شديد فإن الخطأ من النوع الأول هو : " رفض فرض صحيح".

الخطأ من النوع الثاني : Type II error

وفي المقابل فإن الخطأ من النوع الثاني يعني " قبول الفرض العدمي بينما هو خاطئ " أي أنه على الرغم من أن الفرض العدمي خاطئ وكان من الواجب رفضه فقد تم أخذ قرار خاطئ بقبوله وباختصار شديد فإن الخطأ من النوع الثاني هو " قبول فرض خاطئ ".

وقد يتساءل البعض عند مدى إمكانية تصغير الخطأين معاً ولكن لسوء الحظ لا يمكن تصغيرهما معاً إلى أدنى حد ممكن، ويبدو أن الطريقة الوحيدة المتاحة لذلك هي زيادة (أو تكبير) حجم العينة، الأمر الذي قد لا يكون ممكنا في كل الحالات. لذلك فإن الذي يحدث عادة هو تثبيت أحدهما كأن يكون نسبة أو احتمال حدوث الخطأ من النوع الأول ومحاولة تصغير الآخر.

مثال (١): -

عينة عشوائية حجمها 49 شخصاً اختيرت من أفراد دولة ما، فإذا كان الوسط الحسابي لدخول الأفراد الأسبوعية في العينة هو 75 ريال . كيف يمكن اختبار الفرض الصفري بأن متوسط الدخل الأسبوعي لمواطني هذه الدولة يساوي 72 ريال مقابل الفرض البديل أنه لا يساوي 72 وذلك بمستوى معنوية % 5 إذا علمت أن الانحراف المعياري لدخول الأفراد يساوي 14 ريال .

الحل : -

متوسط العينة

متوسط المجتمع

١- الفرض العدمي : هو أن متوسط المجتمع يساوي 72 وبالرموز: $Ho : \mu = 72$

٢- الفرض البديل : هو أن المتوسط لا يساوي 72 وبالرموز: $H1 : \mu = 72$

٣- الإحصائية: بما أن العينة كبيرة فإن الإحصائية في حالة اختبار الوسط تأخذ الشكل التالي:

$$Z_{\bar{x}} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{75 - 72}{\frac{14}{\sqrt{49}}} = 1.5$$

حيث $n = 49 , \sigma = 14 , \bar{X} = 75 , \mu = 72$

وبالتعويض نحصل على :

$$Z_{\bar{x}} = \frac{75 - 72}{\frac{14}{\sqrt{49}}}$$

$$Z_{\bar{x}} = \frac{3}{\frac{14}{7}} = \frac{3}{2} = 1.5$$

أي أن قيمة الإحصائية تساوي 1.5

رفض

قبول H0

رفض

1.96-

1.96+

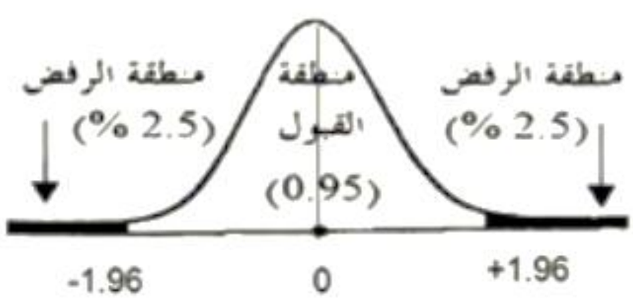
تابع الحل : -

٤- حدود منطقتي القبول والرفض: نحصل عليها من التوزيع الطبيعي المعياري حيث مستوى المعنوية 5% وبما أن الفرض البديل هو: "لا يساوي" فإن ما يستخدم في هذه الحالة هو اختبار الطرفين كما في الشكل التالي :

α = ٥٪

ثقة ٩٥٪

z = ١.٩٦ جدولية



تابع الحل : -

وقد حصلنا على حدود منطقتي القبول والرفض وذلك بقسمة درجة الثقة (المكملة لمستوى المعنوية) والتي تساوي 0.95 على 2 فنحصل على 0.4750 وبالكشف في جدول التوزيع الطبيعي المعياري عن Z التي تقابل المساحة 0.4750 نجد أنها تساوي 1.96 وحيث أنها موزعة على طرفي المنحنى بالتساوي فنضع إشارة موجبة في النصف الأيمن، وإشارة سالبة في النصف الأيسر، أي أن منطقة القبول تبدأ من القيمة ١.٩٦- وتستمر حتى القيمة ١.٩٦ + (أي أن أي قيمة محصورة بين هاتين القيمتين تكون في منطقة القبول، وأي قيمة خارج هذه الحدود تكون في منطقة الرفض).

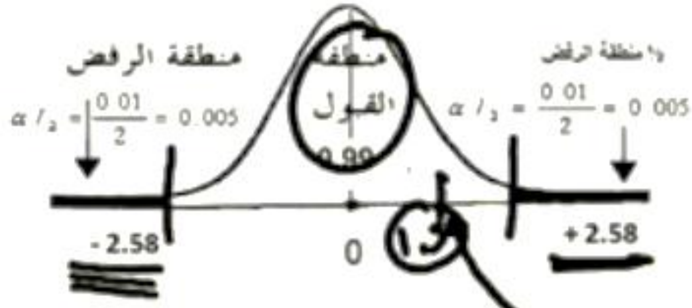
Z = ١,٥

٥- المقارنة والقرار: وبمقارنة قيمة الإحصائية المحسوبة من الخطوة رقم 3 (والتي تساوي 1.5) بحدود منطقتي القبول والرفض (من الخطوة رقم 4) نجد أنها تقع في منطقة القبول ولذلك فإن القرار هو:

قبول الفرض الصفري بأن متوسط دخول الأفراد الأسبوعية في هذه الدولة يساوي 72 دولارا وذلك بمستوى معنوية 5 %.

تابع الحل : -

لو استخدمنا مستوى معنوية 1% بدلاً من 5% كما في المثال أعلاه فإن حدود منطقتي القبول والرفض تصبح كما يلي :



ثقة %99

Z = 1.5

وبمقارنة قيمة الإحصائية 1.5 بحدود منطقتي القبول والرفض نجد أنها تقع في منطقة القبول أي أن القرار هو نفسه قبول الفرض الصفري ولن يتغير بل يتأكد باستخدام مستوى معنوية 1%.

مثال (٤) :-

$n=49$

عينة عشوائية حجمها 49 شخصاً اختيرت من أفراد دولة ما، فإذا كان الوسط الحسابي لدخول الأفراد الأسبوعية في العينة هو 75 $\bar{X}=75$ ريال . كيف يمكن اختبار الفرض العدمي بأن متوسط الدخل الأسبوعي لمواطني هذه الدولة يساوي 72 $\mu=72$ ريال مقابل الفرض البديل أنه لا يساوي 72 وذلك بمستوى معنوية 5 % إذا علمت أن الانحراف المعياري لدخول الأفراد يساوي 14 ريال

$$Z=\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}=\frac{75-72}{\frac{14}{\sqrt{49}}}=1.5$$

$H_0: \mu = 72$

$H_1: \mu \neq 72$

طرفين

الحل :-

١- الفرض العدمي : هو أن متوسط المجتمع يساوي 72 وبالرموز :

$Ho : \mu = 72$

٢- الفرض البديل : هو أن المتوسط لا يساوي 72 وبالرموز :

$Ho : \mu \neq 72$

٣- الإحصائية : بما أن العينة كبيرة فإن الإحصائية في حالة اختبار الوسط تأخذ الشكل التالي :

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

حيث أن :-

$n = 49, \sigma = 14, \bar{X} = 75, \mu = 72$

تابع الحل :-

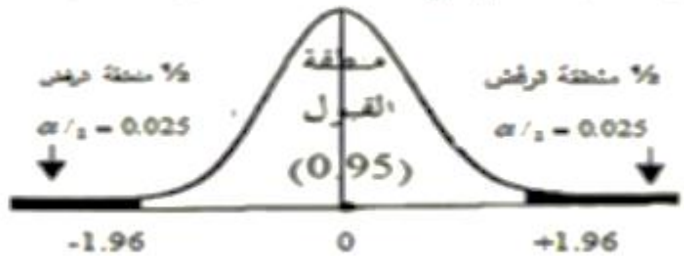
وبالتعويض نحصل على :-

$$z_{\bar{x}} = \frac{75-72}{\frac{14}{\sqrt{49}}}$$

$$z_{\bar{x}} = \frac{3}{\frac{4}{7}} = \frac{3}{2} = 1.5$$

أي أن قيمة الاحصائي تساوي 1.5.

٤- حدود منطقتي القبول و الرفض : نحصل عليها من التوزيع الطبيعي المعياري حيث مستوى المعنوية 5% وبما أن الفرض البديل هو : "لا يساوي " فإن ما يستخدم في هذه الحالة هو اختبار الطرفين كما في الشكل التالي :-



تابع الحل :-

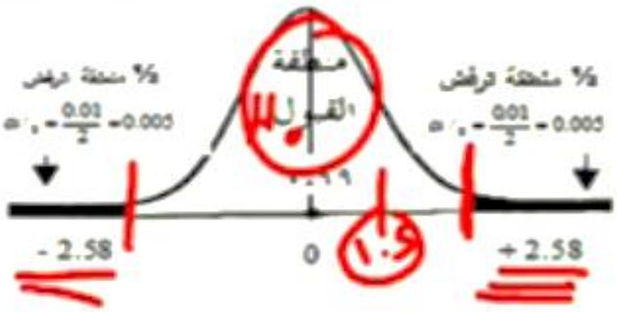
٥- المقارنة والقرار : وبمقارنة قيمة الإحصائية المحسوبة من الخطوة رقم 3 (والتي تساوي 1.5) بحدود منطقتي القبول والرفض (من الخطوة رقم 4) نجد أنها تقع في منطقة القبول لذلك فإن القرار هو :

قبول الفرض العدمي بأن متوسط دخول الأفراد الأسبوعية في هذه الدولة يساوي 72 دولاراً وذلك بمستوى معنوية 5 %.

تابع الحل :-

ملاحظة :

لو استخدمنا مستوى معنوية 1% بدلاً من 5% كما في المثال أعلاه فإن حدود منطقتي القبول والرفض تصبح كما يلي :



وبمقارنة قيمة الإحصائية 1.5 بحدود منطقتي القبول والرفض نجد أنها تقع في منطقة القبول أي أن القرار هو نفسه قبول الفرض العدمي ولن يتغير بل يتأكد باستخدام مستوى معنوية 1%.

مثال (٥)

$H_0 \; P = 0.70$

$H_1 \; P < 0.70$

يدّعى أحد المرشحين في الانتخابات أنه سيحصل على نسبة 70% من أصوات الناخبين عندما تجري الانتخابات. ولاختبار هذا الادعاء تم اختيار عينة عشوائية من الناخبين حجمها 100 ناخب، ووجد أن نسبة من يؤيدون المرشح في العينة هي % 60 اختبر مدى صحة ادعاء المرشح بأن النسبة في المجتمع هي % 70 مقابل الفرض البديل أن النسبة أقل من 70% وذلك بمستوى معنوية % 5.

$P = 70\%$ $n = 100$ $\hat{P}$

الحل :-

١- الفرض العدمي هو أن النسبة في المجتمع (نسبة من يؤيدون المرشح في المجتمع) هي 0.70 أي أن الفرض العدمي هو أن الادعاء صحيح وأن المرشح سيحصل على النسبة التي ادعاها وهي % 70 بالرموز HO : P = 0.70

٢- الفرض البديل والمنطقي : في هذه الحالة هو أن النسبة في المجتمع أقل من هذا الادعاء وبالرموز : H1 : P < 0.70

تابع الحل

$$Z = \frac{\hat{P} - P}{\sqrt{P(1-P)}} = \frac{0.60 - 0.70}{\sqrt{0.70 \times 0.30}} = -2.17$$

٣- الإحصائية : وتأخذ الإحصائية في حالة اختبار النسبة الشكل التالي :

$$Z_{\hat{p}} = \frac{\hat{P} - P}{\sqrt{\frac{P(1-P)}{n}}}$$

حيث أن :-

$$n = 100, \hat{P} = 0.60, P = 0.70 \cdot 1 - p = 1 - 0.70 = 0.30$$

$$Z_{\hat{p}} = \frac{0.60 - 0.70}{\sqrt{\frac{0.70 \times 0.30}{100}}}$$

$$= \frac{-0.10}{0.046}$$

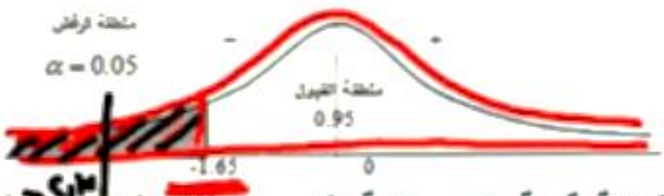
$$Z_{\hat{p}} = -2.17$$

أي أن قيمة الإحصائية تساوي 2.17 -

تابع الحل :-

٩٥٪
معنوية ٥٪ بطرف واحد
ح الجدولية ١٫٦٥

٤- حدود منطقتي القبول والرفض نحصل عليها من التوزيع الطبيعي المعياري، حيث مستوى المعنوية $\alpha = 5\%$ وبما أن الفرض البديل هو " أقل من " فنستخدم اختبار الطرف الأيسر.



٥- المقارنة والقرار : وبمقارنة قيمة الإحصائية التي حصلنا عليها في الخطوة رقم (٣) التي تساوي 2.17 - بحدود منطقتي القبول والرفض (من الخطوة رقم ٤) نجد أن قيمة الإحصائية تقع في منطقة الرفض لأن 2.17 - أصغر من 1.65 - فإن القرار هو :

<u>رفض الفرض العدمي بادعاء المرشح بأن نسبة مؤيديه في المجتمع هي 70 % وقبول الفرض البديل بأن النسبة أقل من 70 % وذلك بمستوى معنوية 5 % (أي أن احتمال الخطأ في هذا القرار لا يتعدى 5 %).</u>

مثال (٦):-

البيانات التالية تمثل نتائج عينتين عشوائيتين مستقلتين مسحوبتين من منطقتين لمقارنة متوسط عمر الناخب فيهما :

حيث $\bar{X}_1 = 35, \bar{X}_2 = 29, n_2 = 80, n_1 = 100$

اختبر الفرض العدمي : أن متوسط عمر الناخب في المنطقة الأولى يساوي متوسط عمر الناخب في المنطقة الثانية بمستوى معنوية 5% مقابل الفرض البديل أنهما غير متساويين إذا علمت أن :

$\sigma_1^2 = 60, \sigma_2^2 = 32$

$H_0 \quad \mu_1 = \mu_2$

$H_1 \quad \mu_1 \neq \mu_2$

الحل :-

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

١- الفرض العدمي أن المتوسطين متساويان وبالرموز :

$$Ho: \mu_1 = \mu_2$$

٢- الفرض البديل أن المتوسطين غير متساويين وبالرموز :

$$HI: \mu_1 \neq \mu_2$$

٣- الإحصائية : تأخذ الشكل التالي :

$$Z_{\bar{X}1-\bar{X}2} = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} =$$

تابع الحل :-

وبالتعويض عن :- $n_1 = 100, n_2 = 80, \bar{X}_1 = 35, \bar{X}_2 = 29, \sigma_1^2 = 60, \sigma_2^2 = 32$

نحصل على :-

$$Z_{\bar{X}1-\bar{X}2} = \frac{35 - 29}{\sqrt{\frac{60}{100} + \frac{32}{80}}}$$

$$= \frac{60}{\sqrt{0.60 + 0.40}}$$

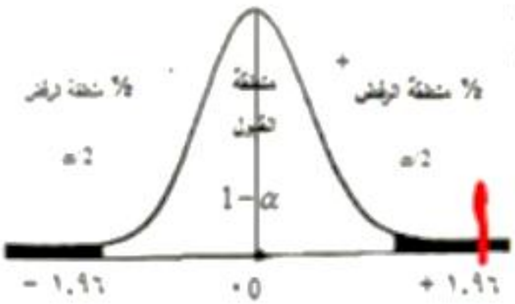
$$= \frac{6}{\sqrt{1}} = 6$$

أي أن قيمة إحصائي الاختبار تساوي 6 .

رفض الفرض العدمي

تابع الحل :-

٤- حدود منطقتي القبول والرفض التي نحصل عليها من جدول التوزيع الطبيعي Z لأن العينات كبيرة، والاختبار هو اختبار الطرفين (لأن الفرض البديل لا يساوي)
ومستوى المعنوية المطلوب هو % 5.



أي أن منطقة القبول تبدأ من 1.96- إلى 1.96+ ومنطقة الرفض هي القيم التي أصغر من 1.96 - والتي أكبر من 1.96 +.

تابع الحل :-

٥- المقارنة والقرار ولما كانت قيمة الإحصائية (والتي تساوي) 6 تقع في منطقة الرفض فإن القرار هو رفض الفرض العدمي وقبول الفرض البديل بمستوى معنوية 5% أي أننا نرفض الفرض القائل بأن متوسط عمر الناخب في المنطقة الأولى يساوي متوسط عمر الناخب في المنطقة الثانية وذلك بمستوى معنوية % 5.

مثال على اختبار Z : هام

إذا كان متوسط استهلاك الفرد السعودي من الدجاج حسب تقارير وزارة الصحة هو (١٢) $\mu=12$ كيلوجرام بانحراف معياري (٦) $\sigma=6$ كيلوجرامات لفترة السبعينات الميلادية. أجرى أحد الباحثين دراسة في عام ٢٠٠٣م من عينة قوامها (٤٩) $n=49$ فرداً ووجد أن متوسط الاستهلاك للفرد هو (١٤) $\bar{X}=14$ كيلوجرام. هل تشير الدراسة الحالية أن متوسط الاستهلاك أرتفع عما عليه في السبعينات.

لو لم يحدد مستوى المعنوية 5%

درجة الثقة 95%

الحل:

$H_0 \quad \mu=12$

$H_1 \quad \mu>12$

$$Z=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}=\frac{14-12}{6/\sqrt{49}}=2.33$$

١) فرض العدم والفرض البديل.

فرض العدم: H_0: μ=12

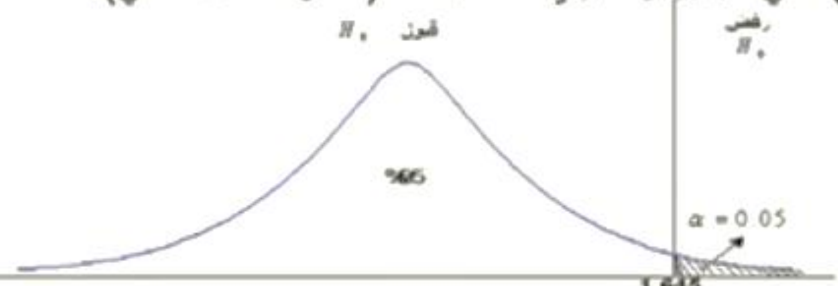
الفرض البديل: H_1: μ>12

٢) مستوى الدلالة =(0.05):

٣) إحصائية الاختبار (Z):

$$Z=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}=\frac{14-12}{6/\sqrt{49}}=2.33$$

٤) تحديد قيمة Z المعيارية من الجدول عند مستوى دلالة (٠.٠٥)، نحتاج لتحديد قيمة Zα التي تقع علي اليمين وتساوي ١.٦٤٥ (أنظر الشكل التالي):



١,٦٥ $Z = ٢,٣٣$

٥) بما أن القيمة المحسوبة أكبر من القيمة النظرية المستخرجة من الجدول كما يبين الشكل، فإنها تقع في منطقة الرفض. وبذلك نرفض فرض العدم حيث أن البيانات المتوفرة تقدم دليلاً كافياً على أن متوسط استهلاك الفرد من لحوم الدواجن في الوقت الحالي قد ارتفع بمستوى معنوي أو ذو دلالة عما علية في سبعينات القرن الماضي.

مثال على اختبار t :

هـ

$n = 250$

لو كانت لدينا عينة عشوائية تتكون من ٢٥٠ طالب وجد أن الوسط الحسابي لأطوال طلاب العينة ١٥٥.٩٥سم، والانحراف المعياري = ٢.٩٤ سم، علما بأن الوسط الحسابي لأطوال طلاب الجامعة يبلغ ١٥٨ سم، اختبر أهمية الفرق المعنوي بين الوسط الحسابي لأطوال طلاب العينة والوسط الحسابي لأطوال طلاب الجامعة .

$\bar{X} =$ S M_0

$t =$

الحل :

سيتم اختبار الفرضيات التالية :

الفرضية الصفرية : لا توجد فروق ذات دلالة إحصائية بين متوسط أطوال الطلاب في العينة ومتوسط أطوال الطلاب في الجامعة

$(\mu = \mu_0)$

الفرضية البديلة : توجد فروق ذات دلالة إحصائية بين متوسط أطوال الطلاب في العينة ومتوسط أطوال الطلاب في الجامعة

$(\mu \neq \mu_0)$

مستوى الدلالة : α = ٠.٠٥

منطقة الرفض : قيمة (ت) الجدولية عند مستوى دلالة α ٠.٠٥ ودرجات حرية ٢٤٩ = $\boxed{1.960}$

n - 1

249=250-1

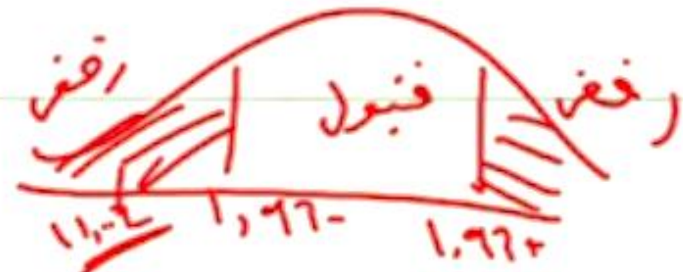
المختبر الإحصائي :

$\bar{X}$ = ١٥٥,٩٥سم ، n = ٢٥٠ طالب ، S= ٢,٩٤ سم

μ = ١٥٨ سم

$t = \frac{155.95-158}{2.94/\sqrt{250}} = -11.006$

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}} = \frac{155.95 - 158}{2.94/\sqrt{250}} = -11.006$$



القرار :

• قيمة ت المحسوبة (- ١١.٠٠٦) أكبر من قيمة ت المجدولـة (١.٩٦) عند مستوى دلالة α = ٠.٠٥ .

∴ نرفض الفرضية الصفرية ونقبل البديلة .
أي أنه توجد فروق ذات دلالة إحصائية بين الوسط الحسابي للعينة والوسط الحسابي لمجتمـع البحث .

✓أنقر بعد ذلك على زر "موافق" OK سيؤدي ذلك إلى تنفيذ الاختبار، وستلاحظ ظهور النتائج في شاشة المخرجات كالتالي :

→ T-Test

$t = -11.006$

درجة الحرية

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
الطول	250	[illegible]	2.9422	.1861

One-Sample Test

	Test Value = 158					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference Lower	Upper
الطول	-11.006	249	.000	-2.0480	-2.4145	-1.6815

قيمة إحصائي الاختبار

0.05 > 0.000

Sig

يتضح من النتائج أن قيمة (ت) المحسوبة t-test = -١١.٠٠٦ ، ودرجات الحرية df = ٢٤٩ ، وقيمة Sig. (2-tailed) =٠.٠٠٠ ، وبما أن قيمة Sig. (2-tailed) في الجدول (٠.٠٠٠) أصغر من قيمة α = ٠.٠٥ فإننا بالتالي نرفض الفرضية الصفرية، أي أنه توجد فروق ذات دلالة إحصائية بين متوسط أطوال العينة ومتوسط أطوال طلاب الجامعة.

نقبل الزعم البديل

Sig < 0.05 البديل

Sig < 0.05
نقبل H_1

✓أنقر بعد ذلك على زر "موافق" OK سيؤدي ذلك إلى تنفيذ الاختبار، وستلاحظ ظهور النتائج في شاشة المخرجات كالتالي :

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	POSTTEST	68.6000	100	8.0000	.8000
	PRETEST	64.2000	100	7.0000	.7001

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	POSTTEST & PRETEST	100	.458	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	POSTTEST - PRETEST	4.3800	7.8579	.7857	2.8210	5.9390	5.575	99	.000

الزعم البديل

أصحاب الفرضية

t = 5.575

df = 99

Sig = 0.000

أقل من 0.05

H_1 نقبل

H_0

نلاحظ أن برنامج الـ SPSS قام مباشرة بحساب الإحصاءات الأساسية للبيانات مثل المتوسط الحسابي للمتغير Posttest (٥٨.٦٦٠) والانحراف المعياري لنفس المتغير (٨.٠٠) ، أما المتغير Pretest فقد كان المتوسط الحسابي (٥٤.٢٨٠) والانحراف المعياري (٧.٠٠) . بالإضافة إلى ذلك تم حساب معامل ارتباط بيرسون للمتغيرات موضع الدراسة Paired Sample Correlation وقد كانت قيمته (٠.٤٥٨) .

ثم بعد ذلك قام البرنامج بحساب قيمة (ت) للمتغيرات موضع الدراسة في الجدول المعنون بـ "اختبار العينات المرتبطة" Paired Sample Test ، ومن هذه النتائج نلاحظ أن قيمة (ت) المحسوبة t-test = ٥.٥٧٥ ، ودرجات الحرية df = ٩٩ ، وقيمة Sig. (2-tailed) =٠.٠٠٠ ، وبما أن قيمة Sig. (2-tailed) في الجدول (٠.٠٠٠) أصغر من قيمة α = ٠.٠٥ فإننا بالتالي نرفض الفرضية الصفرية ، أي أن أداء الطلاب في الكلية بعد أخذ البرنامج التدريبي كان أفضل من أدائهم قبل أخذ البرنامج التدريبي عند مستوى α = ٠.٠٥ .

مثال (٢) : -

قام أحد الباحثين بتفريغ ما تم الحصول عليه من معلومات في جدول تحليل التباين كالتالي :

مصدر التباين	مجموع المربعات SS	درجات الحرية df	متوسط المربعات Means	قيمة F
بين المجموعات Between groups	200	10	20	٤
داخل المجموعات Within groups	50	10	5	
الكلي (المجموع) Total	250	20		

تابع المثال :-

قيمة إحصائي الاختبار F تساوي :-

(أ) 10

(ب) 5

(ج) 80

(د) لا شيء مما سبق

٤

من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض (إذا علمت أن قيمة F الجدولية تساوي 6.88) يمكن :-

٦,٨٨ > ٤

(أ) قبول الفرض البديل .

(ب) قبول الفرض العدمي

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق

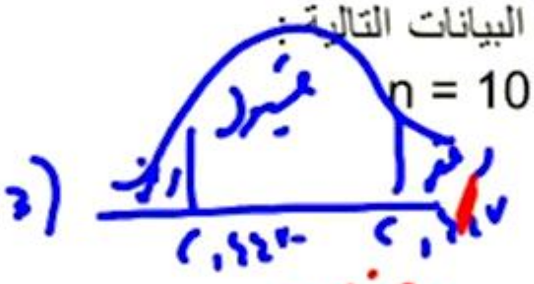
هـ

مثال:

اختبر معنوية معامل الارتباط لوكان لدينا البيانات التالية :

$r = 0.91$ ، $n = 10$

وذلك بمستوى معنوية 5 %.



1) $H_0 : R = 0$

$H_1 : R \neq 0$

2) $T = \dfrac{r}{\sqrt{\dfrac{1-r^2}{n-2}}} = \dfrac{0.91}{\sqrt{\dfrac{1-0.91^2}{10-2}}} = 6.2$

3) رفض الفرض العدمي وقبول الفرض البديل

الحل :

لوكان لدينا البيانات التالية :

$r = 0.91$ ، $n = 10$

وتكون خطوات اختبار معنوية الارتباط كما يلي :

<u>١ – الفرض العدمي :</u> $H_0 : R = 0$

<u>٢ – الفرض البديل :</u> $H_A : R \neq 0$

٣ – الإحصائية :

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.91}{\sqrt{\frac{1-(0.91)^2}{10-2}}} = \frac{0.91}{\sqrt{\frac{0.1719}{8}}} = \frac{0.91}{\sqrt{0.0215}} = \frac{0.91}{0.1466} = 6.208$$

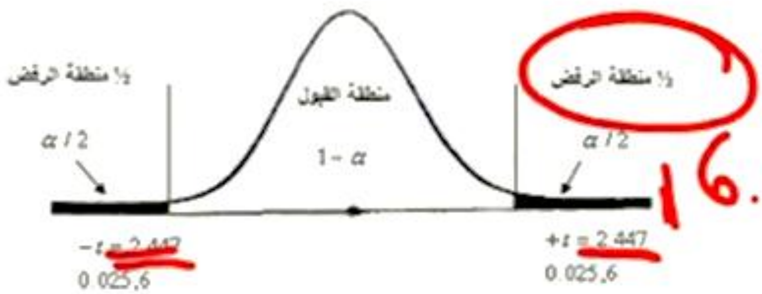
إذا : $t = 6.208$

٤ – حدود منطقتي القبول والرفض:

من جدول t حيث مستوى المعنوية $\alpha = 0.05$ ، $\frac{\alpha}{2} = 0.025$ ودرجات الحرية تساوي ($n - 2 = 10 - 2 = 8$) نجد أن قيمة t تساوي ٢,٤٤٧ وتكون حدود منطقتي القبول والرفض كما يلي:

عمادة التعلم الإلكتروني والتعليم عن بعد
Deanship of E-Learning and Distance Education

جامعة الملك فيصل
King Faisal University



٥ – **<u>المقارنة والقرار :</u>** بمقارنة قيمة الإحصائية المحسوبة في الخطوة رقم ٣ والتي تساوي 6.2074 بحدود منطقتي القبول والرفض (أو قيم t الجدولية في الخطوة رقم ٤) نجد أنها تقع في منطقة الرفض (حيث أنها أكبر من 2.447) لذلك فإن القرار هو : **<u>رفض الفرض العدمي وقبول الفرض البديل</u>. أي رفض أن معامل الارتباط يساوي صفر.** وقبول أن معامل الارتباط **<u>لا</u>** يساوي صفر أي **<u>يوجد ارتباط</u>** بين المتغيرين (أعمار الناخبين و دخولهم اليومية) وذلك بمستوى معنوية 5 %.

مثال :-

" أن معامل الإرتباط بين ثلاث ظواهر إقتصادية قد بلغت ($r = 0.21$) و كان عدد المفردات التي تم دراستها ($n = 10$)، وقد رغب الباحث في دراسة معنوية الارتباط و ذلك بمستوى 5% "

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.21}{\sqrt{\frac{1-0.21^2}{10-2}}}$$

(1) قيمة إحصائي الاختبار t تساوي :-

(أ) 0.6075

(ب) -0.6075

(ج) 6.2

(د) لا شيء مما سبق

الجدولي 2.447

(٢) إذا علمت أن حدود منطقتي القبول و الرفض هي (2,447- , 2.447)فعلى ذلك يمكن :-

-2.447   2.447

(أ) قبول الفرض العدمي

(ب) رفض الفرض العدمي .

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق

تمرين واجب :-

إذا علمت أنه :-

" أن معامل الإرتباط بين ثلاث ظواهر إقتصادية قد بلغت ($r = 0.91$) و كان عدد المفردات التي تم دراستها ($n = 10$)، وقد رغب الباحث في دراسة معنوية الارتباط و ذلك بمستوى 5% "

(١) قيمة إحصائي الاختبار t تساوي :-

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}} = \frac{0.91}{\sqrt{\frac{1-0.91^2}{10-2}}}$$

(أ) 0.6208

(ب) -0.6208

(ج) 6.208 > 2.447

(د) لا شيء مما سبق

(٢) إذا علمت أن حدود منطقتي القبول و الرفض هي (2,447- , 2.447)فعلى ذلك يمكن :-

(أ) قبول الفرض العدمي.

(ب) رفض الفرض العدمي

نقبل الفرض البديل

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق

مثال :-

$$r_{1.2.3}=\frac{r_{1.2}-(r_{2.3}\cdot r_{1.3})}{\sqrt{(1-(r_{1.3})^2)(1-(r_{2.3})^2)}}$$

يقوم أحد الباحثين بدراسة العلاقة بين ثلاث من الظواهر و هي A و B و C ، وجد أن الارتباط بين كل من الظاهرة الأولى A و الظاهرة الثانية B يساوي (0.62) و الارتباط بين الظاهرة الأولى و الثالثة يساوي (-0.225) و الارتباط بين كل من الظاهرة الثانية و الثالثة تساوي (0.179) ، فالمطلوب تقدير قيمة الارتباط الجزئي بين كل من هذه الظواهر .

$r_{1.2}=0.62$

$r_{1.3}=-0.225$

$r_{2.3}=0.179$

$$r_{1.2.3}=\frac{(620)-[(-.225)(179)]}{\sqrt{[1-(-.225)^2][1-(179)^2]}}$$

$$=\frac{(620)+(0402)}{\sqrt{(1-0.0506)(1-0.032)}}$$

$$=\frac{0.662}{\sqrt{(0.9494)(0.968)}}=\frac{0.662}{\sqrt{0.919}}$$

$$=\frac{0.660}{0.9586}=0.689$$

مثال (٣) :-

قامت إحدى شركات الادوية بتوريد ١٠٠ كرتونه مصل الحمة الشوكية لأحد المستشفيات كل كرتونة تحتوى على ٣٠ زجاجة مصل و لوحظ توزيع عدد زجاجات المصل المكسورة بالكرتونة وكان كما يلي :-

عدد الزجاجات المكسورة بالكرتونه	0	1	2	3	4	5	المجموع
عدد الكراتين	22	28	35	10	3	2	100

و المطلوب : توفيق دالة إحتمال توزيع ذات الحدين لعدد زجاجات المصل المكسورة بالكرتونة في الشركة واختبار جودة التوفيق عند درجة الثقة 95% .

الحل :-

دالة التوزيع الاحتمال لتوزيع ذات الحدين تتوقف على معلمتين n و p أي عدد الفئات و الاحتمال :-

أولاً عدد الفئات تساوى 5 أي أن n = 5 .

ثانياً : الاحتمال :-

نحسب المتوسط أولا

$$\mu = \frac{0\times22+1\times28+2\times35+3\times10+4\times3+5\times2}{100} = 1.5$$

لا تنسى أن :-

$$\mu = n\,p$$

$$1.5 = 5 \times P$$

$$P = 0.3$$

الحل :-

H_0 : عدد زجاجات المصل المكسورة بالكرتونة الواحدة يتبع التوزيع ثنائي الحدين بالمعلمتين $n = 5 , p = 0.3$.

H_1 : عدد زجاجات المصل المكسورة بالكرتونة الواحدة لا يتبع التوزيع ثنائي الحدين بالمعلمتين $n = 5 , p = 0.3$.

و من خلال الاعتماد على معلمات التوزيع ثنائي الحدين يمكن تكوين جدول توزيع ثنائي الحدين ، كما يتضح من الجدول التالي :-

الحل :- $^nC_r (0.3)^0 \times (0.7)^5$

عدد الزجاجات المكسورة	الاحتمال	التكرار المتوقع
0	$^5c_0 \times (0.3)^0 \times (0.7)^5 = 0.1681$	16.81
1	$^5c_1 \times (0.3)^1 \times (0.7)^4 = 0.3602$	36.02
2	$^5c_2 \times (0.3)^2 \times (0.7)^3 = 0.3087$	30.87
3	$^5c_3 \times (0.3)^3 \times (0.7)^2 = 0.1323$	13.23
4	$^5c_4 \times (0.3)^4 \times (0.7)^1 = 0.0284$	2.84
5	$^5c_5 \times (0.3)^5 \times (0.7)^0 = 0.0024$	0.24
المجموع	**1**	**100**

$\boxed{100}$

لاحظ أن التكرار المتوقع = الاحتمال × عدد الكراتين ١٠٠

تابع الحل :-

ولان إختبار كا٢ يشترط ألا يقل التكرار المتوقع لاي خلية عن ٥ ، لذلك سيتم دمج الخلايا الثلاثة الاخيرة لكي يصبح التكرار المتوقع لهم معا أكبر من أو يساوي ٥ كما يتضح من الجدول التالي :-

عدد الزجاجات المكسورة	التكرارات المشاهدة ش	التكرارات المتوقعة ت	(ش – ت)٢	(ش – ت)٢ / ت
0	22	16.81	26.94	1.60
1	28	36.02	64.32	1.79
2	35	30.87	17.06	0.55
3-5	15	16.31	1.72	0.11
المجموع	100	100		4.05

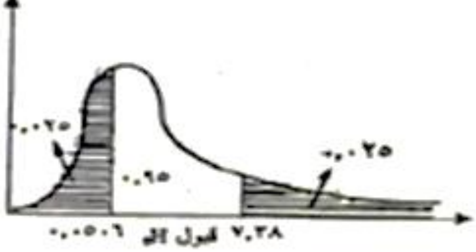
إذاً كا ٢ المحسوبة = 4.05

تابع الحل :-

درجات الحرية = عدد الخلايا بعد الدمج – عدد المعلمات =

= ٤ – ٢ = ٢

و عند مستوى معنوية 5% فإن قيمتي كا ٢ الجدولية هما (7.38 , 0.0506) و تكون منطقتي القبول و الرفض للفرض العدمي كما يلي :-



و حيث أن قيمة كا ٢ المحسوبة تقع في منطقة القبول لذلك نقبل الفرض العدمي و هو ما يعني أن منحنى التوزيع ثنائي الحدين يعتبر توفيق جيد لتوزيع عدد الزجاجات حسب الزجاجات المكسورة .

مثال (٦) :-

قام أحد الباحثين باختبار مدى أتفاق نتائج الطلاب للمعدلات التراكمية مع التوزيع المنتظم و حصل على النتائج التالية :-

٤٠٠

فئات المعدل التراكمي	التكرارات المشاهدة ش	التكرارات المتوقعة ت
0 -	56	80
1 -	70	80
2 -	106	80
3 -	90	80
4 - 5	78	80
المجموع	400	400

تابع مثال (٦) :-

المطلوب :-

١- تقدير قيمة كا٢ المحسوبة .

٢- إذا علمت أن حدود قيمة كا٢ الجدولية هي **(11.1 , 0.484)** فهل يمكن قبول الفرض العدمي .

تابع الحل :-

(ش – ت)٢ / ت	(ش – ت)٢	التكرارات المتوقعة ت	التكرارات المشاهدة ش	فئات المعدل التراكمي
7.2	576	80	56	0 -
1.25	100	80	70	1 -
8.45	676	80	106	2 -
1.25	100	80	90	3 -
0.05	4	80	78	4 - 5
18.2		400	400	المجموع

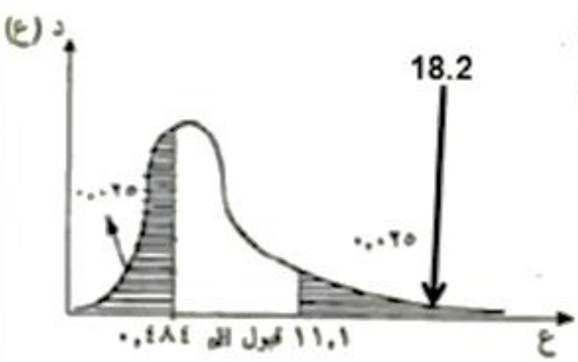
إذاً كا ٢ المحسوبة = 18.2

تابع الحل :-

درجات الحرية = (عدد الفئات – ١)

= (٥-١) = 4

و عند مستوى معنوية 5% فإن قيمتي كا ٢ الجدولية هما (11.1 , 0.484) و تكون منطقتي القبول و الرفض للفرض العدمي كما يلي :-



و حيث أن قيمة كا ٢ المحسوبة تقع في منطقة الرفض.

لذلك نقبل الفرض البديل و هو ما يعني أن منحنى التوزيع المنتظم يعتبر توفيق جيد لتوزيع طلاب الكلية حسب فئات المعدل التراكمي .

تابع الحل :-

الجدول الثالث يبين نتيجة اختبار مربع كاي كالتالي:

(1-2) (1-5)

قيمة الاختبار

درجة الحرية

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2.437[a]	4	.656
Likelihood Ratio	2.459	4	.652
Linear-by-Linear Association	.298	1	.585
N of Valid Cases	72		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 5.35.

مستوى دلالة الاختبار

0.05

الأدبي

يبين الجدول الثالث السابق أن قيمة اختبار مربع كاي هي ٢.٤٣٧ بدرجة حرية مقدارها ٤

يبين لنا من الجدول أن أقل قيمة لمستوى الدلالة هي ٠.٦٥٦ = Asymp. Sig. (2-sided) وهي اكبر من مستوى الدلالة ٠.٠٠٥ = α وبالتالي لا نستطيع رفض الفرضية الصفرية أي أن تقدير الطالب لا يعتمد على جنسه.

H_0 ✓

مثال (٦) :-

الجدول التالي يوضح نتيجة إختبار مربع كاي (كا٢) عند مستوى معنوية 5% :-

Chi-Square Test

	Value	df	Asymp. Sig (2-sided)
Pearson Chi-Square	1.9496	3	.0437
Likelihood Ratio	1.9672	3	.0434
Linear-by-Linear Association	.2384	1	.0390
N of Valid Cases	32		

< 0.05

أجب عن الاسئلة التالية من خلال النتائج الواردة في الجدول السابق :-

0.0437 - 0.05

أصغر = −

أكبر = +

الحل :-

(١) قيمة إحصائي الاختبار كا٢ تساوي :-

(أ) .2384

(ب) 1.9672

(ج) 1.9496

(د) لا شيء مما سبق

(٢) قيمة مستوى الدلاله المحسوبة للإختبار تساوي :-

(أ) .0437

(ب) .0434

(ج) .0390

(د) لا شيء مما سبق

(٣) من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض يمكن :-

(أ) قبول الفرض البديل

(ب) قبول الفرض العدمي .

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق

تابع الحل :-

Ranks المتوسط

	CODES	N	Mean Rank	Sum of Ranks
SAMPLES	2	10	11.10	111.00
	3	10	9.90	99.00
	Total	20		

Test Statistics[b]

مان ويتني

	SAMPLES
Mann-Whitney U	44.000
Wilcoxon W	99.000
Z	-.457
Asymp. Sig. (2-tailed)	.648
Exact Sig. [2*(1-tailed Sig.)]	.684[a]

> 0.05

تابع الحل :-

يلاحظ من نتائج هذا الإختبار: أن قيمة P.Value تساوى 0.648 وهي أكبر من مستوى المعنوية ٥% وبالتالى فاننا نقبل الفرض العدمي بأن متوسط درجات مادة المحاسبة فى كلية إدارة الأعمال جامعة الملك فيصل يساوى متوسط درجات مادة المحاسبة فى جامعة الدمام اى ان الفروق بين الجامعتين غير معنوية.

مثال (٢) :-

" قام أحد الباحثين بمقارنة عينة من مرتبات موظفي القطاع الحكومي من مدينة الرياض بأخرى من مدينة جدة وذلك بصدد الوقوف على ما إذا كان هناك إختلاف في متوسط المرتبات وذلك عند مستوى معنوية 5%، وبإستخدام البرنامج الاحصائي SPSS حصلنا على النتائج التالية :-

Test Statistics

	SAMPLES
Mann-Whitney U	55.000
Wilcoxon W	95.000
Z	-.037
Asymp. Sig. (2-tailed)	.028
Exact Sig. [2*(1-tailed Sig.)]	.034

< 0.05

الحل :-

(١) الاختبار المستخدم لدارسة الفرق بين متوسطي مجتمعين في هذه الحالة :-

(أ) كا٢ .

(ب) مان وتني

(ج) ويلكوكسون .

(د) لا شيء مما سبق

(٢) قيمة إحصائي الاختبار تساوي :-

(أ) -.037

(ب) .028

(ج) .034

(د) لا شيء مما سبق

(٣) من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض يمكن :-

(أ) قبول الفرض البديل

(ب) قبول الفرض العدمي

(ج) عدم قبول أي من الفرضين .

(د) لا شيء مما سبق .

تابع الحل :-

هام

Ranks

		N	Mean Rank	Sum of Ranks
AFTER - BEFORE	Negative Ranks	7[a]	4.93	34.50
	Positive Ranks	1[b]	1.50	1.50
	Ties	0[c]		
	Total	8		

Test Statistics[b]

	AFTER - BEFORE
Z	-2.313[a]
Asymp. Sig. (2-tailed)	.021

< 0.05

تابع الحل :-

قام البرنامج بحساب الفروق في الوزن على أساس التالي:

الفرق = الوزن بعد ممارسة الرياضة – الوزن قبل ممارسة الرياضة

ويلاحظ أيضا: أن متوسط الرتب السالبة (٤.٩٣) أكبر من متوسط الرتب الموجبة (١.٥) وهذا معناه أن متوسط الوزن قبل ممارسة الرياضة أكبر من متوسط الوزن بعد ممارسة الرياضة (إذا فى غاية الأهمية أن نعرف الترتيب الذى إستخدمه البرنامج للعينتين)

تابع الحل :-

ويلاحظ من نتائج هذا الاختبار أن قيمة P.Value تساوي 0.021 وهي أقل من مستوى المعنوية 5% وبالتالي فإننا نقبل الفرض البديل بأن متوسط الوزن قبل ممارسة الرياضة يختلف معنويًا عن متوسط الوزن بعد ممارسة الرياضة.

تابع المثال :-

Ranks

		N	Mean Rank	Sum of Ranks
AFTER-BEFORE	Negative Ranks	7	2.36	43.50
	Positive Ranks	1	3.54	3.54
	Ties	0		
	Total	8		

Test Statistics

	AFTER-BEFORE
Z	-.313
Asymp. Sig. (2-tailed)	.421

> 0.65

الحل :-

(١) من الجداول السابقة يمكن توضيح أن :-

(أ) مستوى الطلاب قبل الحصول على البرنامج التدريبي أفضل من المستوى بعد الحصول على البرنامج .
(ب) مستوى الطلاب بعد الحصول على البرنامج التدريبي أفضل من المستوى قبل الحصول على البرنامج .
(ج) مستوى الطلاب قبل الحصول على البرنامج التدريبي مساوي لمستوى بعد الحصول على البرنامج .
(د) لا شيء مما سبق

(٢) من خلال مقارنة قيمة إحصائي الاختبار بقيمة حدود منطقتي القبول والرفض يمكن :-

(أ) قبول الفرض البديل .
(ب) قبول الفرض العدمي .
(ج) عدم قبول أي من الفرضين .
(د) لا شيء مما سبق .

تابع الحل :-

Ranks

	CODES	N	Mean Rank
SAMPLES	1	8	16.88
	2	8	10.75
	3	8	9.88
	Total	24	

Test Statistics[a,b]

	SAMPLES
Chi-Square	4.706
df	2
Asymp. Sig.	.095

0, > 0.05

مثال :-

" قام أحد الباحثين بدراسة درجات مجموعة من الطلاب في مادة التحليل الاحصائي في ثلاث جامعات هي: جامعة الملك فيصل – جامعة الدمام – جامعة الملك سعود ، وذلك لدراسة مدى وجود إختلاف بين مستوى الطلاب في الجامعات الثلاثة السابقة بإستخدام إختبار كروسكال- والس وذلك عند مستوى معنوية 5%، تم الحصول على النتائج التالية بإستخدام البرنامج الاحصائي SPSS:-

Test Statistics

	SAMPLES
Ci-Square	.706
df	2
Asymp. Sig.	.025

$H_0: M_1 = M_2 = M_3$

$H_1: M_1 \neq M_2 \neq M_3$

$.025 < 0.05$

(١) من الجدول السابق يمكن :-

(أ) قبول الفرض البديل القائل بمعنوية الفروق بين الجامعات الثلاثة .

(ب) قبول الفرض العدمي القائل بأن الفروق بين الجامعات الثلاثة غير معنوية .

(ج) قبول الفرض العدمي القائل بأن الفروق بين الجامعات الثلاثة معنوية .

(د) لا شيء مما سبق .

NPar Tests

One-Sample Kolmogorov-Smirnov Test

		Dinner	
N		50	حجم العينة
Normal Parameters a,b	Mean	15.26	متوسط البيانات
	Std. Deviation	6.782	الانحراف المعياري للبيانات
Most Extreme Differences	Absolute	.081	أكبر فرق بين البيانات ودالة التوزيع الاحتمالية
	Positive	.081	
	Negative	-.069	
Kolmogorov-Smirnov Z		.573	قيمة اختبار جودة المطابقة
Asymp. Sig. (2-tailed)		.898	مستوى دلالة الاختبار

a. Test distribution is Normal.

b. Calculated from data.

تبين النتائج أعلاه أن متوسط عدد الزبائن هو ١٥.٢٦ بانحراف معياري قدره ٦.٧٨٢ وأن قيمة اختبار كولموجروف سميرنوف لجودة المطابقة هو ٠.٥٧٣.

NPar Tests

One-Sample Kolmogorov-Smirnov Test

		Dinner	
N		50	حجم العينة
Normal Parameters[a,b]	Mean	15.26	متوسط البيانات
	Std. Deviation	6.782	الانحراف المعياري للبيانات
Most Extreme Differences	Absolute	.081	اكبر فرق بين البيانات ودالة التوزيع الاحتمالية
	Positive	.081	
	Negative	-.069	
Kolmogorov-Smirnov Z		.573	قيمة اختبار جودة المطابقة
Asymp. Sig. (2-tailed)		.898	مستوى دلالة الاختبار

a. Test distribution is Normal.

b. Calculated from data.

طبيعي

> 0.05

تبين النتائج أعلاه أن متوسط عدد الزبائن هو ١٥.٢٦ بانحراف معياري قدره ٦.٧٨٢ وأن قيمة اختبار كولموجروف سميرنوف لجودة المطابقة هي ٠.٥٧٣

القرار:

يبين الجدول السابق أن قيمة مستوى دلالة الاختبار هي٠.٨٩٨ = *Asymp. Sig. (2-tailed)* وهي اكبر من مستوى دلالة الفرضية الصفرية ٠.٠٥ = α وبالتالي نقبل الفرضية الصفرية، أي أن البيانات تتبع التوزيع الطبيعي وبالتالي نستنتج ان البيانات تتبع التوزيع الطبيعي بمتوسط قدره 15.26 وانحراف معياري 6.782 أي $X: N(15.26, 6.782)$

وإذا أردنا اختبار أن التوزيع يتبع توزيع بواسون نختار من الشاشة المخصصة لذلك توزيع بواسون وهكذا مع باقي التوزيعات.