مصور أفريقيا

الاحصاء

Sheet (1)

د.عادل وحيدة

[مع تمنياتنا للجميع بالتوفيق والنجاح]

الاسم: ....

خريف -2022

الإحتمالات

مصطلح الاحتمال: ويعني مقدار ثقتنا في إمكانية حدوث
شيء غير مؤكد الوقوع وهذا المفهوم
للإحتمال لا يوجد له تفسير علمي واحد متفق عليه
فهناك من يفسره بأنه التكرار النسبي لحدوث نتيجة
عندما تتكرر التجربة تحت ظروف متشابهة ويسمى
بالتعريف التجريبي لمصطلح الاحتمال وهناك تفسير
تقليدي مبني على أساس مفهوم النتائج ذات الفرص
المتساوية مثل إلقاء قطعة نقود متزنة مرة واحدة
سيكون هناك نتيجتان إما H & T وعليه فكلا
الاحتمالين للصورة أو الكتابة هو 50% أي $\frac{1}{2}$
وهناك تفسير ذاتي مبني على أن الشخص يعطي احتمالا
معينا لنتيجة ما في تجربة معينة بناءا على المعلومات
التي لديه ومدى خبرته، وعليه فإننا سنعتمد على
النظرية الرياضية للإحصاء والاحتمالات والتي تلعب دورا
هاما في الاختبارات والقرارات.

بعض المصطلحات المتعلقة بالإحتمالات

1- **التجربة العشوائية**: هي التجربة التي تكون نتائجها غير
معروفة مسبقا بشكل أكيد، ومن أمثلتها:

1. إلقاء قطعة نقود إذ إننا نعلم أنه لها نتيجتين (H, T)
ولكن لا نستطيع تحديد أي منهما مسبقا.
2. ملاحظة وحدة منتجة في أحد المصانع من حيث إنها
صالحة أو غير صالحة (G, D)
3. عدد المكالمات التي يستقبلها أحد المكاتب في ساعة معينة
هو 0,1,2,3

2- **فراغ العينة**: S أو Ω

وهو مجموعة جميع النتائج الممكنة للتجربة العشوائية ويرمز لها
بالرمز S أو Ω.

مثال: فراغ العينة لقطعة نقود القيت مرتين هو:

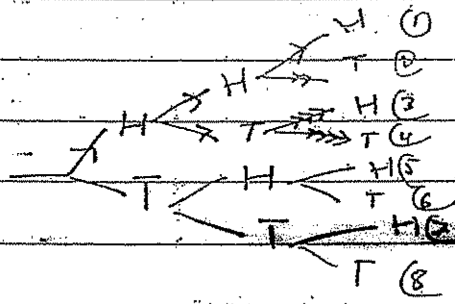
$$S = \{HH, HT, TH, TT\}$$

ويمكن عناصرها بصفة عامة عند القائها n من المرات
هي $n(S) = 2^n$

فمثلا عندما $n=3$ يكون عدد عناصر فراغ العينة $n(S)=2^3=8$

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

ولتسهيل الحصول عليها نتبع طريقة الشجرة كما هو مبين



مثال: أكتب فراغ العينة لتجربة عشوائية متمثلة في إلقاء قطعة
نقود معدنية حتى ظهور الصورة لأول مرة.

الحل

$$S = \{H, TH, TTH, TTTH, ----\}$$

وهي مجموعة غير محدودة النتائج ذات العناصر لانهائية

3 **الحدث**: هو مجموعة جزئية من فراغ العينة S ويرمز
للأحداث عادة بالحروف الكبيرة A, B, C, ...
فإذا احتوى الحدث على عنصر واحد سمي حدثا بسيطاً،
أما إذا احتوى على أكثر من عنصر من عناصر S فيسمى
حدثا مركبًا.

مثال / أخذت عينة عشوائية مكونة من ثلاثة مصابيح كهربائية وتم فحصها بغرض معرفة صلاحيتها من عدمه أوجد:

1) فراغ العينة للنتائج الممكنة S

2) عناصر الحدث A الذي يمثل أن عدد المصابيح الغير صالحة D في هذه العينة لا يزيد عن 1 ؟

3) كل المصابيح صالحة في حدث B

الحل

$S = \{GGG, GGD, GDG, GDD, DDG, DGD, DGG, DDD\}$

$A = \{DGG, GDG, GGD, GGG\}$

$B = \{GGG\}$

ملاحظات مهمة: على الأقل 2 مثلا يعني $x \geq 2 \Rightarrow \{2,3,...\}$

على الأكثر 2 مثلا يعني $x \leq 2 \Rightarrow \{0,1,2\}$

أكثر من 2 مثلا يعني $x > 2 \Rightarrow \{3,4,5,...\}$

أقل من 2 مثلا يعني $x < 2 \Rightarrow \{0,1\}$

4 الحدث المكمل: يعرف الحدث المكمل للحدث A مثلا بأنه الحدث الذي يحتوي على جميع عناصر S التي لا يحتويها الحدث A ويرمز له بالرمز $\acute{A}$ أو $\bar{A}$

ومنه نستنتج أن $A \cup \acute{A} = S$ & $A \cap \acute{A} = \phi$

5 الحدث المؤكد: هو الحدث الذي يحتوي على جميع عناصر S أي أن $A = S$ ويسمى أيضا الحدث الشامل.

6 الحدث المستحيل: هو الذي لا يحتوي على أي عنصر من عناصر S أي أن $A = \phi$

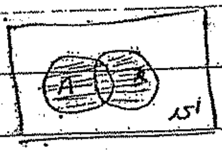
7 الأحداث المتنافية: وهي الأحداث المانعة لبعضها البعض فيقال أن A&B متنافيان إذا كان حدوث أحدهما يمنع حدوث الآخر أي أنه يستحيل حدوثهما معا أي أن $A \cap B = \phi$ وكذلك إذا كان $A_1, A_2, ..., A_n$ أحداث متنافية

فإنه $A_i \cap A_j = \phi$ لكل $i \neq j$ حيث $i,j = 1,2,\ldots,n$ (أي لا يوجد مشترك بين أي اثنين منهم)

[8] الأحداث المستقلة : هي الأحداث التي لا تؤثر في بعضها البعض فحدوث أحدها لا يؤثر ولا يتأثر بحدوث أو عدم حدوث الأحداث الأخرى.

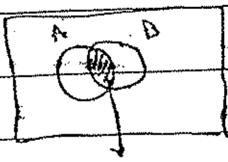
ج/ العمليات على الأحداث

الاتحاد Union ($\cup$) : ويعبر اتحاد الحدثين A & B عن وقوع أحدهما على الأقل بمعنى وقوع A أو B أو كلاهما ويعبر عنه رياضيا ($A \cup B$) وتكون عناصر الاتحاد من عناصر الحدثين معاً دون تكرار في حالة الاستفادة بشكل فن كما هو موضح



$A \cup B$

التقاطع ($\cap$) Intersection : ويعبر عن وقوع الحدثين معاً في آن واحد ويشمل كل النتائج المشتركة بين الحدثين ويعبر عنه رياضيا ($A \cap B$) كما في الشكل



$A \cap B$

[9] الأحداث الشاملة : يقال للحدثين A & B بأنهما شاملان إذا كان $A \cup B = S$

متباينين فإن A & A^c يعتبران حدثين شاملين لأن $A \cup A^c = S$

نلاحظ أن:

$(A^c)^c = A \qquad \phi^c = S$

$A^c = S - A \quad \& \quad A \cap S = A$

$A \cup S = S \quad \& \quad A \cap \phi = \phi$

$A \subset B \Rightarrow A \cap B = A \quad \& \quad A \cup B = B$

$(A\cup B)^c = A^c \cap B^c$ & $A\cap B^c = A - A\cap B$

$(A\cap B)^c = A^c \cup B^c$ $B\cap A^c = B - A\cap B$

$A\cup B = A + B - A\cap B$

الاحتمال: هو دالة P تحدد قيمة تعبر عن فرصة الظهور لكل حدث ينتمي لفراغ العينة، بحيث تحقق الخصائص الآتية:

* لأي حدث A ينتمي إلى Ω يكون
* $P(\Omega)=1$ & $P(\phi)=0$
* $P(A^c) = 1 - P(A)$
* إذا كان $A \subseteq B$ $\therefore P(A) \leq P(B)$
* وإذا كان الحدث A عناصر عددها $n(A)$ فإن $\boxed{P(A) = \frac{n(A)}{n(\Omega)}}$
* لأي حدث A فإن $\boxed{0 \leq P(A) \leq 1}$

مثال: ألقيت زهرة نرد منتظمة واحدة أوجد احتمال ظهور عدد يقبل القسمة على 3 ؟

الحل

$\Omega = \{1,2,3,4,5,6\} \Rightarrow n(\Omega) = 6$

نفرض أن الحدث A هو ظهور عدد يقبل القسمة على 3

$\therefore A = \{3,6\} \Rightarrow n(A) = 2$

$P(A) = \frac{n(A)}{n(\Omega)} = \frac{2}{6} = \boxed{\frac{1}{3}}$

مثال: عند رمي قطعة نقود ثلاث مرات أوجد احتمال الحصول على صورتين على الأقل

الحل

$n(\Omega) = 2^3 = 8$

نفرض أن A = يمثل الحدث الحصول على صورتين على الأقل:

$A = \{HHT, HTH, THH, HHH\}$

$\therefore n(A) = 4 \Rightarrow P(A) = \frac{4}{8} = \boxed{0.5}$

الدرس 29

طرق عد عناصر فراغ العينة

1/ **قاعدة الجمع**: إذا كان للتجربة A عدد (n_1) من النتائج للتجربة B (n_2) من النتائج، وكانت التجربتان متنافيتين، فإن عدد النتائج للتجربة A أو B هو $\boxed{n_1 + n_2}$

ويمكن تعميمها على مجموعة من التجارب المتنافية، فإن عدد الطرق لحدوث أي من هذه التجارب هو $\boxed{n_1 + n_2 + n_3 + \cdots + n_k}$

مثال/ تتنافس مجموعة من الشركات على أربعة مشاريع من المجموعة A وخمسة مشاريع من المجموعة B، بحيث لا يحق لأي شركة أن تأخذ أكثر من مشروع واحد فما هو عدد الطرق التي يمكن بها أن تأخذ شركة معينة مشروعًا من هذه المشاريع؟

الحل

∵ الشركة لا يمكنها إلا مشروعًا واحدًا فقط فإن هذا المشروع إما أحد مشاريع المجموعة A أو أحد مشاريع المجموعة B

∴ عدد الطرق = 4 + 5 = 9 طرق

2/ **قاعدة الضرب**:

تجربة A ⟵ n_1 من الطرق

تجربة B ⟵ n_2 من الطرق

وكانت التجربتان غير متنافيتين (أي يمكن وقوعهما معًا)، فإن عدد الطرق لوقوعهما معًا هو $n_1 \times n_2$

على التعميم لهذه القاعدة:

A B C ---- K

n_1 n_2 n_3 n_k

فإن عدد عناصر فراغ العينة لهذه التجارب معًا

$$\boxed{n(S) = n_1 \times n_2 \times n_3 \times \cdots \times n_k}$$

مثال: ما عدد عناصر فراغ العينة $n(S)$ في الحالات الآتية:

1) إلقاء قطعة نقود وحجر نرد معًا؟

2) أربع نقود معًا؟ (3) ثلاث زهرات نرد معًا؟

١- الحل

قطعة نقود: H T

$n_1 = 2$

زهرة النرد: 1 2 3 4 5 6

$n_2 = 6$

∴ عدد الطرق لإلقائهما معاً = $2 \times 6 = 12$ والهذه الطريقة

$S = \{H1, H2, \ldots, H6, T1, T2, \ldots, T6\}$

٢- اربعة نقود ∴ $n(S) = 2 \times 2 \times 2 \times 2 = 16$

٣- ثلاث زهرات ∴ $n(S) = 6 \times 6 \times 6 = 216$

3. قاعدة التوافيق: وهي عدد الطرق التي من الممكن لها أن نأخذ (r) عنصر من بين (n) من العناصر بحيث لا يكون Combination (للترتيب أهمية) ويسمى n توافيق r ويرمز لها بالرمز $^{n}c_{r}$ حيث $r < n$ وتعرف كالآتي:

$$\boxed{^{n}c_{r} = \frac{n!}{r!(n-r)!}}$$

مثال/ ما عدد الطرق التي تكفي لاختيار 3 مهندسين لتكوين مجلس استشاري هندسي من بين ثمانية مهندسين؟

الحل

لاحظ أن الترتيب غير مهم لأن اختيار المهندسين ABC مثلاً هو نفسه BAC أو CAB، لأن كل منهم عضو في المجلس.

$n(S) = {}^{8}c_{3} = \frac{8!}{3!5!} = \frac{8 \times 7 \times 6 \times \cancel{5!}}{3 \times 2 \times \cancel{5!}}$

$n(S) = 56$

ملاحظة: إذا سحب «r» عنصر من «n» من العناصر بالإرجاع أي أن العنصر المسحوب يعاد مرة أخرى

قبل سحب العنصر الذي يليه فإن $n(S) = n^r$.

(4) **قاعدة التباديل: Permutation**

التباديل: هي ترتيب لعدة أشياء بأخذها كلها أو بعضها في كل مرة مع (مراعاة الترتيب) فعدد التباديل لمجموعة مكونة من (n) من الأشياء مأخوذة (r) في كل مرة يساوي عدد الترتيبات المختلفة الممكن تكوينها من (n) من الأشياء بحيث تحتوي كل ترتيبة على (r) من الأشياء المختلفة ويحسب كالتالي

$$\boxed{{}^{n}P_{r} = \frac{n!}{(n-r)!}} \qquad r \le n$$

لاحظ أن $\quad {}^{n}P_{r} = n!$

مثال/ بكم طريقة يمكن أن تحصل سبع شركات على ثلاثة عقود مختلفة بحيث لا يمكن لأي شركة الفوز بأكثر من عقد

الحل

∵ العقود مختلفة فإن الحدث $ع_1ع_2ع_3 \neq ع_3ع_2ع_1$

وعليه فإن للترتيب أهمية ∴ عدد العناصر الكلية هو

$$n(S) = {}^{7}P_{3} = \frac{7!}{4!} = \frac{7\times6\times5\times\cancel{4!}}{\cancel{4!}} = 210$$

نوع المعاينة

- بدون إعادة
  - بدون مراعاة الترتيب: ${}^{n}C_{r}$ ✓
  - مراعاة الترتيب: ${}^{n}P_{r}$ ✓
- مع الإعادة
  - بدون مراعاة الترتيب: ${}^{n+r-1}C_{r}$
  - مراعاة الترتيب: n^{r}

5 قاعدة التقسيم :- هي عدد الطرق التي علينا بها تقسيم n من العناصر إلى (K) مجموعة بحيث تحتوي كل مجموعة على $n_1, n_2, n_3, \dots, n_K$ على التوالي من العناصر المختلفة

$$n(S) = \frac{n!}{n_1!\, n_2! \dots n_K!}$$

مثال ما عدد الطرق التي يتم بها توزيع 15 مهندساً على مجموعة من الشركات بحيث تحتاج الشركة الأولى إلى ثلاثة مهندسين والثانية إلى أربعة مهندسين والثالثة إلى ستة مهندسين والرابعة إلى اثنين مهندسين؟

الحل

$n = 15 \quad n_1 = 3 \quad n_2 = 4 \quad n_3 = 6 \quad n_4 = 2$

$$n(S) = \frac{15!}{3!\,4!\,6!\,2!} = 6,306,300 \text{ طريقة}$$

مثال بكم طريقة علينا اختيار رئيس ونائبه وسكرتيره من بين 7 مهندسين لتشكيل مجلس إدارة هندسي؟

الحل

بـ المناصب مختلفة :. عدد الطرق الممكنة لتشكيل المجلس هي

$^7P_3 = 210$ ((لأن الترتيب مهم))

مثلا $ABC \neq CBA$

س ١ ، س ٢

نريد أن نشكل مجلساً هندسياً من 4 مهندسين من بين 3 مهندسين كهرباء و 4 مهندسين مدني و 3 مهندسين ميكانيكيين بحيث

أ) يكون بالمجلس 2 كهربائيين وواحد مدني وآخر ميكانيكي ؟

ب) لا يكون بالمجلس مهندساً مدنياً ؟

ج) يكون بالمجلس مهندساً مدنياً ؟

قانون الجمع

$p(A\cup B') = p(A) + p(B') - p(A\cap B')$

$= p(A) + p(B') - [p(A) - p(A\cap B)]$

$= p(B') + p(A\cap B)$

مسلمات الاحتمالات

1- لأي حدث A يكون $0 \le P(A) \le 1$

2- $P(\phi)=0$ & $P(S)=1$

3- احتمال حدوث A أو B أو كليهما (بمعنى على الأقل أحدهما) يعرف بأنه

$P(A\cup B) = P(A) + P(B) - P(A\cap B)$

وعليه التعميم

$P(A\cup B\cup C) = P(A)+P(B)+P(C)-P(A\cap B)$
$-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C)$

4- وإذا كان A & B متنافيان فإن $P(A\cup B) = P(A)+P(B)$

5- = = = = مستقلان فإن $P(A\cup B)=P(A)+P(B)-[P(A)P(B)]$

6- احتمال حدوث A دون B أو بمعنى A فقط يعرف بأنه

$P(A\cap B^c) = P(A) - P(A\cap B)$ ملاحظة $p(A)=p(A\cap B')+p(A\cap B)$

7- احتمال حدوث B دون A أو بمعنى B فقط هو

$P(B\cap A^c) = P(B) - P(A\cap B)$

8- احتمال عدم حدوث A وعدم حدوث B: $P(A^c\cap B^c) = P(A\cup B)^c = 1-P(A\cup B)$

9- احتمال عدم حدوث أحدهما على الأقل: $P(A^c\cup B^c) = P(A\cap B)^c = 1-P(A\cap B)$

مثال/ اخترت عينة عشوائية من طلبة الإحصاء مجموعها 60 طالب، فوجد أن 15 طالب مسجلين في المقرر ST305 & 12 طالب مسجلين في مقرر ST401 & 6 مسجلين في المقررين معاً. فإذا تم اختيار أحد الطلبة عشوائياً من هذه العينة أوجد:

أ) احتمال أن يكون الطالب مسجل في أحد المقررين على الأقل؟

ب) = = = غير مسجل في أي من المقررين؟

ج) = = = مسجل في المقرر ST305 فقط؟

الحل

نفرض أن الحدث A هو أن الطالب مسجل في المقرر ST305:

$P(A) = 15/60$

والحدث B هو أن الطالب مسجل في المقرر ST401:

$P(B) = 12/60$

وبذلك يكون $P(A\cap B) = 6/60$

1. يكون المطلوب الأول هو $P(A\cup B) = P(A)+P(B)-P(A\cap B)$

$$= \frac{15}{60}+\frac{12}{60}-\frac{6}{60} = \frac{21}{60}$$

2. احتمال عدم تسجيله في أي من المقررين هو: $P(A^c\cap B^c)$

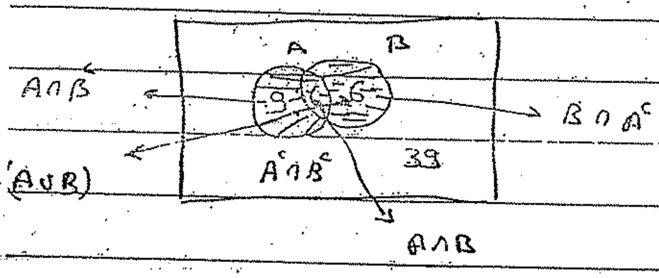
والذي يساوي $= 1-P(A\cup B)$

$$= 1-\frac{21}{60} = 39/60$$

3. احتمال تسجيله في المقرر ST305 فقط هو: $P(A\cap B^c) = P(A)-P(A\cap B)$

$$= \frac{15}{60}-\frac{6}{60} = 9/60$$

ويمكن توضيح ذلك بالرسم



تمرين

مثال: إذا كان A, B حدثان في تجربة وكان $P(A\cup B) = 3/4$

$P(A) = 1/2 \quad P(B^c) = 5/8$

أوجد (1) $P(A\cap B)$ (2) $P(B\cap A^c)$ (3) $P(A^c\cup B^c)$

الحل

1) $P(A\cup B) = P(A)+P(B)-P(A\cap B)$

$\frac{3}{4} = \frac{1}{2}+\frac{3}{8}-P(A\cap B)$

$\therefore P(A\cap B) = \frac{1}{2}+\frac{3}{8}-\frac{3}{4} = \frac{4}{8}+\frac{3}{8}-\frac{6}{8} = \boxed{\frac{1}{8}}$

(2) $P(B\cap A^c) = P(B)-P(A\cap B) = \frac{3}{8}-\frac{1}{8} = \boxed{\frac{2}{8}}$

3) $P(A^c\cup B^c) = P(A\cap B)^c = 1-P(A\cap B) = 1-\frac{1}{8} = \boxed{7/8}$

مثال: إذا كان A و B حدثان معرفان في تجربة وكان $P(A) = 4/10$ و $P(A\cup B) = 7/10$

أوجد قيمة $P(B)$ في الحالات الآتية (1) A و B متنافيان

ج) A & B مستقلان ؟ الحل

(1) A & B: متنافيان ∴ $P(A \cap B) = 0$ ∴

$P(A \cup B) = P(A) + P(B)$ ∴ $0.7 = 0.4 + w$ ∴ $w = \boxed{0.3}$ #

(2) A & B مستقلان ∴ $P(A \cup B) = P(A) + P(B) - [P(A)\,P(B)]$

∴ $0.7 = 0.4 + w - [0.4\,w]$

$0.3 = 0.6\,w$ ∴ $w = \frac{0.3}{0.6} = \boxed{0.5}$

∴ $\boxed{w = 0.5}$ #

(3) الاحتمال الشرطي

في هذا الباب سندرس الطريقة التي يتغير بها احتمال وقوع الحدث A بعد توفر معلومات عن وقوع حدث آخر وليكن B. هذا الاحتمال الجديد للحدث A يسمى الاحتمال المشروط للحدث A إذا عُلم وقوع الحدث B ويرمز له بالرمز $P(A/B)$ ويقرأ احتمال وقوع A بشرط وقوع B أو احتمال A بمعلومية وقوع B ويعرف كما يلي:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

وعلى ذلك

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

مثال: إذا تم إلقاء زهرة نرد متزنة مرتين ولوحظ أن مجموع الرقمين الظاهرين كان عدد فردي، فما احتمال أن يكون T أقل من 8 ؟

الحل

II \ I	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

مجموع فردي T

مجموع فردي وأقل من 8 T

$\frac{12}{18}$

36 حدث

نفرض أن الحدث B هو أن المجموع T لعدد فردي

A = = = = T أقل من 8

وعليه يكون المطلوب $P(A/B) = \frac{P(A \cap B)}{P(B)}$

$A \cap B = \{(1,2)(2,1)(1,4)(4,1)(2,3)(3,2),(6,1)(1,6)$
$(2,5)(5,2),(4,3),(3,4)\}$

$n(A \cap B) = 12 \qquad n(S) = 36$

$\therefore p(A \cap B) = 12/36 = \boxed{1/3}$

$P(B) = \frac{n(B)}{n(S)} = \frac{18}{36} = \boxed{1/2}$

$\therefore p(A/B) = \frac{1/3}{1/2} = \boxed{\frac{2}{3}}$

مثال / الجدول التالي يبين توزيع الطلبة في كلية العلوم للعام الجامعي (98/99م)

المجموعة / الجنس	العلوم الرياضية A	علوم الحياة B	علوم طبيعية C	المجموع
ذكور M	200	100	100	400
إناث F	400	300	100	800
المجموع	600	400	200	1200

فإذا تم اختيار أحد الطلبة عشوائياً أوجد الاحتمالات الآتية:

1- أن يكون الشخص المختار من العلوم الرياضية وأن يكون ذكراً؟

2- = = = = من علوم الحياة علماً بأنه أنثى؟

3- = = = = = ذكراً بشرط أن يكون من العلوم الطبيعية؟

الحل

(1) $p(A \cap M) = \frac{200}{1200} = \boxed{\frac{1}{6}}$

(2) $p(B/F) = \frac{p(B \cap F)}{p(F)} = \frac{300/1200}{800/1200} = \boxed{3/8}$

(3) $p(M/C) = \frac{p(M \cap C)}{p(C)} = \frac{100/1200}{200/1200} = \boxed{\frac{1}{2}}$

قاعدة الضرب الشرطي

$$P(A\cap B) \rightarrow P(A/B)\,P(B)$$
$$P(A\cap B) \rightarrow P(B/A)\,P(A)$$

نتيجة: إذا كان لدينا حدثين A و B يكون احتمال

$$\boxed{P(A) = P(A/B)P(B) + P(A/B^c)P(B^c)}$$

مثال

ورشة بها 12 سيارة بها عطل بالمكيف فقط و 7 سيارات بها عطل بنظام التبريد للمحرك فقط و 5 سيارات بها عطل بالمكيف وبنظام التبريد للمحرك أخذ أحد الفنيين سيارة عشوائيا فوجد بها عطل بالمكيف فما احتمال أن يكون بها عطل بنظام تبريد المحرك؟

الحل

نفرض أن A هو الحدث الذي يمثل أن السيارة بها عطل بالمكيف

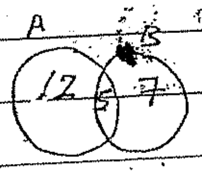
$$P(A) = 12/24$$

ونفرض أن B = = = = = = = عطل بنظام تبريد المحرك:

$$P(B) = 7/24$$

$$P(A\cap B) = 5/24$$

ويكون المطلوب



$$P(B/A) = \frac{P(A\cap B)}{P(A)}$$

$$= \frac{5/24}{12/24} = \boxed{\frac{5}{12}}$$

مثال/ صندوق به 9 كرات بيضاء وأربع كرات حمراء، وصندوق آخر به كرة بيضاء وأخرى حمراء، أخذت كرة واحدة من الصندوق الأول ووضعت في الصندوق الثاني، ثم أخذت كرة من الصندوق الثاني فما احتمال أن تكون الكرة المسحوبة من الصندوق الثاني بيضاء؟

الحل:

نفرض أن B هو الحدث أن الكرة المنقولة من الصندوق الأول بيضاء

∴ B هو الحدث ان الكرة المنقولة من الصندوق الأول حمراء ، وليكن الحدث

A هو أن الكرة المنقولة من الصندوق الثاني بيضاء ، إذن

$P(B) = 2/6 \qquad P(B^c) = 4/6$

I	
2W	1W
4R	1R

$$\therefore P(A) = P(A/B)\,P(B) + P(A/B^c)\,P(B^c)$$

$$= \frac{2}{3} \times 2/6 + \frac{1}{3} \times 4/6 = \frac{8}{18} = \boxed{\frac{4}{9}}$$

أوجد احتمال أن الكرة المسحوبة من الصندوق الثاني حمراء؟

((يترك للطلبة التفكير في ذلك)).

مثال / 40% من الطلبة نجحوا في الإحصاء و 25% نجحوا في الحاسوب و 20% نجحوا في المادتين ، فإذا أختير طالبا عشوائيا أوجد

1 - احتمال أن يكون ناجحا في الحاسوب علما بأنه نجح في الإحصاء؟

2 - = = = = في الإحصاء = = = = الحاسوب؟

3 - = = = = = = = = راسب في الحاسوب؟

4 - = = = راسبا في الحاسوب = = راسب في الإحصاء؟

الحل

$P(S) = 0.4 \qquad P(C) = 0.25 \qquad P(S \cap C) = 0.20$

(1) $P(C/S) = \dfrac{P(C \cap S)}{P(S)} = \dfrac{.20}{.40} = \boxed{0.5}$

(2) $P(S/C) = \dfrac{P(C \cap S)}{P(C)} = \dfrac{.20}{0.25} = \boxed{0.8}$

(3) $P(S/C^c) = \dfrac{P(S \cap C^c)}{P(C^c)} = \dfrac{P(S) - P(S \cap C)}{1 - P(C)}$

$= \dfrac{.40 - .20}{1 - .25} = \dfrac{.20}{.75} = \boxed{0.27}$

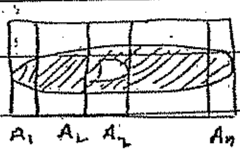
(4) $P(C^c/S^c) = \dfrac{P(C^c \cap S^c)}{P(S^c)} = \dfrac{1 - P(S \cup C)}{1 - P(S)} = \dfrac{0.55}{0.60} = \boxed{0.92}$

نظرية الاحتمال الكلي

إذا كان $A_1, A_2, A_3, \ldots, A_n$ أحداث متنافية والتي من الممكن أن ينتج عن أي منها أو عن كل واحد منها الحدث D، فإن احتمال هذا الحدث D يُحسب

$$\boxed{P(D) = P(D/A_1)P(A_1) + P(D/A_2)P(A_2) + \ldots + P(D/A_n)P(A_n)}$$

البرهان: من الرسم واضح أن



$$D = (A_1 \cap D) + (A_2 \cap D) + \ldots + (A_n \cap D)$$

وباستخدام قاعدة الضرب الشرطي ينتج أن

$$P(D) = P(A_1)P(D/A_1) + P(A_2)P(D/A_2) + \ldots + P(A_n)P(D/A_n) \quad \#$$

مثال/ معمل للتحاليل به ثلاثة فنيين يقوم الأول بتحليل 50% من التحاليل الواردة للمعمل، ويقوم الثاني بتحليل 30%، ويقوم الفني الثالث بتحليل الباقي، فإذا كانت احتمالات الخطأ في التحليل لكل منهم على الترتيب هي 2%، 1%، 4%. فإذا تم اختيار تحليل عشوائياً من هذا المعمل، فما احتمال أن يكون به خطأ في التحليل؟

الحل

$P(F_1) = .50 \qquad P(F_2) = 0.30 \qquad P(F_3) = 0.20$

$P(E/F_1) = 0.02 \qquad P(E/F_2) = 0.01 \qquad P(E/F_3) = 0.40$

المطلوب

$$P(E) = 0.50 \times .02 + 0.30 \times .01 + 0.20 \times .04$$
$$= \boxed{0.021}$$

نظرية بييز: إذا كان الحدث D في نظرية الاحتمال الكلي قد وقع فإننا نريد معرفة أي من الأحداث $A_1, \ldots, A_n$

تابع ٤٠

كان سببًا في وقوعه. فإن احتمال أن يكون بسبب A_1 مثلًا هو:

$$P(A_1/D) = \frac{P(D/A_1)\,P(A_1)}{P(D)}$$

وهكذا بصفة عامة فإن:

$$\boxed{P(A_i/D) = \frac{P(D/A_i)\,P(A_i)}{P(D)}}$$

$i = 1, 2, 3, \ldots, n$

مثال: في المثال السابق: إذا وجد خطأ في التحليل فما احتمال أن يكون من فعل الفني الثاني؟ الحل

أي أن المطلوب هو

$$P(F_2/E) = \frac{P(F_2)\,P(E/F_2)}{P(E)}$$

وفي نفس المثال إذا اختير تحليل ووجد بأنه جيد ما احتمال أن يكون تم به الفني الأول؟ الحل

المطلوب $P(F_1/G)$

$$P(F_1/G) = \frac{P(F_1)\,P(G/F_1)}{P(G)} = \frac{.50 \times .98}{.979} = \boxed{0.501}$$

~~~~

مثال: مكتب هندسي به ٣ سكرتيرات تطبع الأولى ٤٥٪ من رسائل والثانية ٠.٣٥ والثالثة ٠.٢٠ فإذا كانت نسبة الأخطاء المطبعية لكل منهن على التوالي هي: ٣٪ ، ١٪ ، ٢٪ أوجد:

١- احتمال إذا اختار المدير رسالة أن يجد بها خطأ مطبعي؟

٢- وإذا وجد بها خطأ مطبعي، ما احتمال أن تكون من طباعة السكرتيرة الثالثة؟

يترك الحل للطلبة

~~~~

مثال/ إذا كان احتمال أن يكون مصدر الحريق ناتج أحد الحقول النفطية هو A يساوي 40% واحتمال أن يكون مصدره B هو 60% وكان احتمال إطفاء الحريق الذي مصدره A هو 90% بينما احتمال إطفاء حريق من المصدر B هو 45%. فإذا تمكن فريق إطفاء الحريق من إخماد حريق قد حدث فأي المصدرين أكثر احتمالاً بأنه مصدر الحريق؟

الحل

نفرض أن الحدث A: هو أن مصدر الحريق هو A

B = = = = = : B = =

ونفرض أن الحدث F: بأنه قد تم إخماد الحريق

احتمال إخماد أو إطفاء الحريق يساوي:-

$$P(F) = .40 \times .90 + .60 \times .45 = 0.63$$

نريد الآن المقارنة بين احتمال أن يكون بسبب A أي $P(A/F)$ وبين أن يكون بسبب المصدر B أي $P(B/F)$ بتطبيق نظرية بيز يكون

$$P(A/F) = \frac{.90 \times .40}{0.63} = \boxed{0.57}$$

$$P(B/F) = \frac{.45 \times .60}{0.63} = \boxed{0.43}$$

وبالمقارنة يتضح أن المصدر A هو الأكثر احتمالاً أن يكون سبب الحريق بالرغم من أن نسبة حدوثه 40% أقل من المصدر B الذي هو 60%.

المتغيرات العشوائية و توزيعات الإحتمالية

المتغير العشوائي: هو دالة حقيقية تحدد قيمة العددية نتيجة إجراء تجربة عشوائية أي ان المتغير العشوائي يكون معرفاً على فراغ العينة، ويرمز للمتغيرات العشوائية بالحروف الكبيرة مثل X, Y, Z ومن أمثلة ما يلي:

1- المتغير العشوائي X الذي يمثل عدد الصور التي تظهر لأعلى عند القاء قطعة نقود 10 مرات أي أن ($Y = 0, 1, 2, 3, \dots, 10$)

2- الزمن الذي يمر قبل تعطل جهاز كهربائي ($X: x > 0$)

3- عدد الحوادث في تقاطع ما خلال أسبوع ($Z: z = 0, 1, 2, \dots$)

ويوجد نوعان من المتغيرات العشوائية، متغيرات عشوائية منفصلة وأخرى متصلة

المتغيرات العشوائية المنفصلة: يأخذ قيماً منفصلة عن بعضها البعض أي توجد فراغات بين القيم المتتالية بحيث يمكن عد هذه القيم حتى وإن امتدت إلى مالانهاية، مثل م.ع X الذي يمثل عدد السيارات التي ينتجها مصنع خلال شهر أو م.ع Y الذي يمثل عدد مرات انقطاع التيار الكهربائي عن منطقة ما خلال مدة معينة، ففي هذه الحالات يأخذ المتغير العشوائي القيم المنفصلة $0, 1, 2, 3, \dots$، وتجدر الإشارة إلى أن م.ع قد يأخذ قيماً غير صحيحة مثل م.ع Z الذي يمثل أرقام أحذية النساء التي ينتجها مصنع للأحذية $9, 8\frac{1}{2}, 8, 7\frac{1}{2}, 7$

المتغيرات العشوائية المتصلة: وفي هذه الحالة يأخذ قيماً متصلة في فترة معينة وهي قيم لا يمكن عدها، ونتحصل على المتغيرات العشوائية المتصلة في الغالب عن استخدام أحد الأجهزة القياس كقياس الطول و الوزن و الزمن وهكذا.

التوزيعات الإحتمالية المنفصلة: إذا كان X م.ع منفصل فإن توزيعه الإحتمالي يكون منفصلاً وهو عبارة عن قائمة

٢) جدول يبين القيم المختلفة للمتغير العشوائي واحتمال كل قيمة $P(X=x)$
وتسمى دالة كتلة الاحتمال ويرمز لها بالرمز $P(x)$ أو $f(x)$ وهي دالة
موجبة تحقق الشرطين:

① $f(x) = P(X=x) \geq 0$

② $\sum_{\forall x} f(x) = 1$ «شرط دالة احتمالية»

أي إن مجموع الاحتمالات $= 1$

مثال: أوجد التوزيع الاحتمالي لـ م.ع. X الذي يمثل عدد الصور التي تظهر عند رمي قطعة متزنة ٣ مرات؟

الحل

Ω	$X=x$	$P(X=x)$
HHH	$x=3$	$1/8$
HHT	$x=2$	$3/8$
HTH	$x=2$	
HTT	$x=1$	
THH	$x=2$	$3/8$
THT	$x=1$	
TTH	$x=1$	
TTT	$x=0$	$1/8$

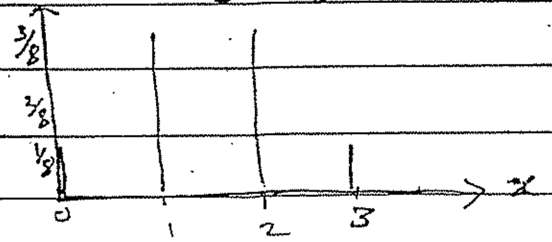
وعليه يكون التوزيع الاحتمالي كما يلي:

$X=x$	0	1	2	3	Σ
$P(X=x)$	$1/8$	$3/8$	$3/8$	$1/8$	1

لاحظ أن هذا التوزيع يمكن كتابته باستخدام دالة الاحتمال كما يلي:

$$f(x) = \begin{cases} \frac{1}{8} & x = 0, 3 \\ 3/8 & x = 1, 2 \end{cases}$$

$f(x) =$



مثال/ إذا كان Y م.ع توزيع كما يلي :-

$Y=y$	-2	-1	0	3	5
$f(y)=p(Y=y)$	0.1	0.3	K	$.2$	$.15$

(i) أوجد قيمة K ؟ ثم أوجد :-

(ii) $p(Y\geq -1)$ (iii) $p(Y<3)$ (iv) $p(Y=2\frac{1}{2})$ (v) $p(Y<3\frac{1}{2})$

الحل

(i) $\because \sum_{\forall y} f(y)=1 \quad \therefore [0.1+0.3+K+0.2+0.15]=1$

$\therefore .75+K=1 \Rightarrow \boxed{K=0.25}$

(ii) $P(Y\geq -1) = P(Y=-1)+P(Y=0)+P(Y=3)+P(Y=5)$

$= 0.3+0.25+0.2+0.15 = \boxed{0.9}$

(iii) $P(Y<3) = P(Y=-2)+P(Y=-1)+P(Y=0)$

$0.1+0.3+.25 = \boxed{0.65}$

(iv) $P(Y=\frac{1}{2}) = 0$

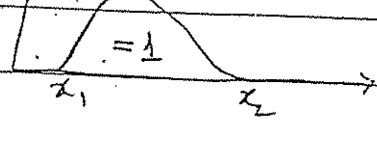
(v) $P(Y<3\frac{1}{2}) = P(Y=3)+P(Y=0)+P(Y=-1)+P(Y=-2)$

$0.2+0.25+0.3+0.1 = \boxed{0.85}$

التوزيعات الاحتمالية المتصلة :- تعرف التوزيع الاحتمالي في هذه الحالة بدالة تسمى دالة كثافة الاحتمال $f(x)$ وتحقق الشرطين التاليين :

(1) $f(x)\geq 0 \quad \forall x$ (2) $\int_{-\infty}^{\infty} f(x)\,dx=1$

فإذا كان X م.ع متصل يأخذ القيم في الفترة $(x_1\ x_2)$ فإن توزيعه الاحتمالي المعرف بالدالة $f(x)$ يمثل بيانيا بمنحنى الدالة $f(x)$ المعرفة في الفترة $(x_1\ x_2)$ والتي تساوي صفراً لجميع القيم خارج هذه الفترة كما في الشكل



(1) المساحة تحت المنحنى تساوي 1

(2) $f(x)$ دالة احتمالية

وتحسب الاحتمالات الخاصة بالمتغير العشوائي المتصل X عن طريق دالة الاحتمال أو دالة الكثافة ، حيث تعرف :-

$$P(a \le x \le b) = \int_a^b f(x)\,dx$$

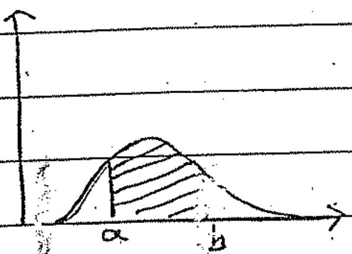
حيث $x_1 \le a < b \le x_2$ ، وحسب هذا التعريف فإنه

$$P(X=a) = \int_a^a f(x)\,dx = 0$$

أي أنه احتمال أن يأخذ م.ع المتصل قيمة معينة a = صفر

وعليه يكون :-

$$P(a \le x \le b) = P(a < x \le b) = P(a \le x < b)$$
$$= P(a < x < b)$$



مثال : إذا كان X م.ع يتوزع بدالة كثافة $f(x)$ حيث

$$f(x) = \begin{cases} \frac{1}{2} & 0 < x < 2 \\ 0 & \text{o.w} \end{cases}$$

(1) أثبت أنها دالة كثافة ؟ ثم أوجد :-

(i) $P(1 \le x \le 2)$ (ii) $P(X < 3/2)$ (iii) $P(X > 1/3)$

(iv) $P(X = 1/2)$

الحل

(1) $\int_{-\infty}^{\infty} f(x)\,dx = \int_0^2 \frac{1}{2}\,dx = \frac{x}{2}\Big]_0^2 = \frac{2-0}{2} = \frac{2}{2} = \boxed{1}$

∴ دالة كثافة $f(x)$:-

(i) $P(1 \le x \le 2) = \int_1^2 \frac{1}{2}\,dx = \frac{x}{2}\Big]_1^2 = \frac{2-1}{2} = \boxed{\frac{1}{2}}$

(ii) $P(X < 3/2) = \int_0^{3/2} \frac{1}{2}\,dx = \frac{x}{2}\Big]_0^{3/2} = \frac{3/2 - 0}{2} = \boxed{\frac{3}{4}}$

(iii) $P(X > \frac{1}{3}) = P(\frac{1}{3} < X < 2) = \int_{1/3}^{2} \frac{1}{2}\,dx = \frac{2-\frac{1}{3}}{2} = \boxed{\frac{5}{6}}$

(iv) $P(X = \frac{1}{2}) = \int_{1/2}^{1/2} \frac{1}{2}\,dx = \frac{x}{2}\Big]_{1/2}^{1/2} = \frac{\frac{1}{2}-\frac{1}{2}}{2} = \frac{0}{2} = \boxed{0}$

مثال إذا كانت
$$f(x) = \begin{cases} c x^2 & 0 < x < 1 \\ 0 & o.w \end{cases}$$

أوجد c ثم أوجد $P(3/8 < x < 5/8)$

الحل

$\because \int_0^1 c x^2\,dx = 1 \Rightarrow c\frac{x^3}{3}\Big]_0^1 = 1$

$= c\left[\frac{1-0}{3}\right] = 1 \Rightarrow \frac{c}{3} = 1 \Rightarrow \boxed{c = 3}$

No 4) $P(3/8 < x < 5/8) = \int_{3/8}^{5/8} 3x^2\,dx$

$= \frac{\cancel{3}x^3}{\cancel{3}}\Big]_{3/8}^{5/8} = \left(\frac{5}{8}\right)^3 - \left(\frac{3}{8}\right)^3 = \frac{125-27}{512} = \boxed{0.19}$

تمارين

1) بين أي من الدوال التالية دالة كتلة

(1) $f(x) = \frac{1}{10}(x-2) \;,\; x = 3,4,5,6$

(2) $f(x) = \frac{1}{20}(2x+4) \;,\; x = -2,-1,0,1,2$

(3) $f(x) = \frac{1}{2}(x-2) \;,\; x = 1,2,3,4$

2) أوجد قيمة c التي تجعل الدالة $f(x)$ دالة كثافة

$f(x) = \frac{3}{16}x^2 \;,\; -c < x < c$

~~$f(x) = 4x^c \;,\; 0 \le x \le 1$~~

القيمة المتوقعة Expected Value

تعريف: القيمة المتوقعة للمتغير العشوائي X هي قيمة متوسط ويرمز لها بالرمز $E(X)$ وتحسب من دالة احتمال كالتالي:

$$E(X) = \begin{cases} \sum x f(x) & \text{if } X \text{ م.ع منفصل Discrete} \\ \int_{-\infty}^{\infty} x f(x)\, dx & \text{if } X \text{ م.ع متصل continuous} \end{cases}$$

خواص التوقع الرياضي:

① $E(c) = c$

② $E(X \pm c) = E(X) \pm c$

③ $E(cX) = cE(X)$

④ $E\left(\sum_{i=1}^{n} X_i\right) = E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n) = \sum_{i=1}^{n} E(X_i)$

⑤ $$Eg(X) = \begin{cases} \sum_{\forall x} g(x) f(x) & X \text{ م.ع منفصل} \\ \int_{-\infty}^{\infty} g(x) f(x)\, dx & X \text{ م.ع متصل} \end{cases}$$

حيث $f(x)$ هي دالة الاحتمال للمتغير العشوائي X.

مثال: اذا كان X م.ع له دالة كتلة معطاة بالجدول التالي: أوجد $E(X)$ & $E(X^2)$؟

الحل:

$X=x$	1	2	3	4	Σ	
$P(X=x)$	$2/5$	$3/10$	$1/5$	$1/10$	1	
$x P(X=x)$	$2/5$	$6/10$	$3/5$	$4/10$	$\frac{20}{10}=2$	$E(X)$
$x^2 P(X=x)$	$2/5$	$12/10$	$9/5$	$16/10$	$\frac{50}{10}=5$	$E(X^2)$

مثال: أوجد قيمة K التي تجعل الدالة $f(x)$ دالة كثافة، ثم أوجد القيمة المتوقعة؟

$$f(x) = \frac{x}{4} + K \qquad 0 \le x \le 2$$

الحل:

$$\int_0^2 \left(\frac{x}{4}+k\right)dx = 1 \quad \therefore \int_0^2 \frac{x}{4}\,dx + \int_0^2 k\,dx = 1$$

$$= \left.\frac{x^2}{8}\right|_0^2 + \left.kx\right|_0^2 = 1$$

$$\frac{4-0}{8} + k(2-0) = 1$$

$$\frac{1}{2} + 2k = 1 \quad \therefore 2k = \frac{1}{2}$$

$$\therefore 4k = 1 \quad \therefore \boxed{k = \frac{1}{4}}$$

$$\therefore f(x) = \frac{x}{4} + \frac{1}{4} = \frac{x+1}{4} \quad 0 \le x \le 2$$

$$\text{Now } E(x) = \int_0^2 x f(x)\,dx = \int_0^2 x\left(\frac{x+1}{4}\right)dx$$

$$= \int_0^2 \frac{x^2}{4} + \int_0^2 \frac{x}{4}\,dx = \left.\frac{x^3}{12}\right|_0^2 + \left.\frac{x^2}{8}\right|_0^2$$

$$= \frac{8-0}{12} + \frac{4-0}{8} = \frac{2}{3} + \frac{1}{2} = \frac{7}{6} = \boxed{1.17}$$

تمرين! (1) اوجد قيمة λ ثم أوجد $E(X)$ ؟

(2 اذا Y م.ع حيث $Y=3X-2$ أوجد $E(Y)$

$X=x$	-1	0	1	2
$P(X=x)$	λ	3λ	0.2	0.4

(3) أوجد قيمة k حيث $P(X \le k) = 0.6$

التباين / وهو مقياس لتشتت القيم وانتشارها حول المتوسط ويرمز له
بالرمز σ^2 أو σ_x^2 ونحسب كما يلي :-

$$V(X) = \sigma^2 = E(X-\mu)^2$$
$$= E(X^2) - \mu^2$$
$$= E(X^2) - [E(X)]^2$$

$\sigma = \sqrt{E(X^2) - [E(X)]^2}$ ∴ الانحراف المعياري σ

خواص التباين :

① $V(K) = 0$

② $V(X \pm K) = V(X)$ ③ $V(KX) = K^2 V(X)$

مثال (12) الخاص بالسابق ص 47 أوجد σ_x^2 ثم أوجد تباين المتغير
$Y_i = 2X - 1$

الحل

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

$\therefore \sigma_x^2 = 5 - [2]^2 = 5 - 4 = \boxed{1}$

Now $V(Y_i) = V(2X-1) = 4V(X) = 4 \times 1 = \boxed{4}$

$\therefore \boxed{\sigma_Y^2 = 4}$ $\therefore \boxed{\sigma_Y = 2}$

مثال أوجد الانحراف المعياري للمتغير العشوائي X الذي دالة كثافته هي :-

$$f(x) \begin{cases} \frac{1}{6}x & 2 < x < 4 \\ 0 & o.w \end{cases}$$

الحل

$$E(X) = \int_2^4 x \cdot \tfrac{1}{6}x \, dx = \frac{1}{6}\frac{x^3}{3}\Big]_2^4 = \frac{56}{18} = \boxed{\frac{28}{9}}$$

$$E(X^2) = \int_2^4 x^2 \tfrac{1}{6}x \, dx = \frac{1}{6}\frac{x^4}{4}\Big]_2^4 = \frac{240}{24} = \boxed{10}$$

$\therefore \sigma_x^2 = 10 - \left[\frac{28}{9}\right]^2 = 0.32$ $\therefore \boxed{\sigma_x = 0.57}$

بعض التوزيعات الاحتمالية الخاصة

توزيع ذي الحدين : Binomial Dist

ويستخدم في الحالات المشابهة للحالة الآتية :

تجربة لها ناتجان فقط يطلق عليهما في العادة "نجاح" و "فشل" حيث تنجح التجربة باحتمال "p" أو تفشل باحتمال "1-p = q" وتعاد التجربة تحت نفس الظروف "أي أن قيمة p ثابتة" ، مع تعاد التجربة n من المرات وكل مرة أو محاولة مستقلة عن الأخرى ويعرف م ع X بأنه يمثل عدد المحاولات الناجحة التي نحصل عليها من عدد "n" من المحاولات المستقلة ، ففي مثل هذه الحالات يتبع م ع X توزيعاً منفصلاً يسمى توزيع ذي الحدين بمعلمتين (n, p) ويرمز له بالرمز $X \sim Bin(n, p)$ ودالة احتمال له بالصورة الآتية

$$\boxed{f(x) = P(X=x) \quad {}^{n}C_{x}\, p^{x} q^{n-x} \qquad x = 0,1,2,3,\ldots,n}$$

بمتوسط قدره $\boxed{\mu = np}$ وتباين قدره $\boxed{V(X) = \sigma^2 = npq}$

ومن أمثلته التطبيقية في الحياة العملية ، وحدات إنتاجية إما صالحة أو غير صالحة ، زرع بذرة إما أن تنبت أو تضمحل ، نتيجة طالب إما نجاح أو رسوب ، جنس مولود إما ذكر أو أنثى ..... الخ

مثال : أخذت عينة عشوائية مكونة من 10 أقلام للتأكد من مطابقتها للجودة ، فإذا كان احتمال أن يكون القلم مطابق للجودة هو 99% أوجد :

1 - احتمال أن تكون جميع الأقلام مطابقة للجودة ؟

2 - = الا يتعدى عدد الأقلام المطابقة للجودة عن 8 أقلام ؟

3 - احتمال أن يكون بالعينة قلم واحد غير مطابق للجودة ؟

4 - إذا اشترى زبون 200 قلم ، ما هو العدد المتوقع للأقلام الجيدة ؟

الحل

حل آخر

نفرض أن X م.ع تمثل عدد الأقلام المطابقة للجودة :.

$X \sim Bin(n, p) \therefore X \sim Bin(10, .99)$

① $P(X=10) = {}^{10}C_{10} (.99)^{10} (.01)^{10-10} = \boxed{0.904}$

② $P(X \leq 8) = 1 - P(X > 8) = 1 - [P(x=9) + P(x=10)]$

$= 1 - [{}^{10}C_{9} (.99)^{9} (.01)^{1} + {}^{10}C_{10} (.99)^{10} (.01)^{10-10}]$

$= 1 - [0.091 + 0.904] = \boxed{0.005}$

③ احتمال قلم واحد غير مطابق

هنا تصبح $p = .01$ & $q = .99$ و يكون المطلوب

$P(X=1) = {}^{10}C_{1} (.01)^{1} (.99)^{9} = \boxed{.091}$

④ $E(X) = np = 200 \times .99 = 198$ قلم

مثال / أطلقت طائرة 6 صواريخ على هدف أرضي، فإذا كان احتمال إصابة أي صاروخ للهدف هو 0.85 ، فأوجد :-

(i) احتمال إصابة جميع الصواريخ الهدف ؟

(ii) احتمال عدم إصابة صاروخين للهدف ؟

(iii) العدد المتوقع لعدد الإصابات ؟

الحل

نفرض أن X م.ع تمثل عدد الصواريخ التي تصيب الهدف

$\therefore X \sim Bin(6, .85)$

ويكون المطلوب (i) هو $P(X=6)$

$\therefore P(X=6) = {}^{6}C_{6} (.85)^{6} (.15)^{0} = 0.377$

(ii) لاحظ أن المطلوب عدم إصابة صاروخين للهدف

هنا يكون م.ع مثلاً Y تمثل عدد الصواريخ التي لم تصب الهدف

تمرين 52

$Y \sim Bin(6, .15)$

وبالتالي تكون المطلوب

$P(Y=2) = {}^{6}C_{2} (.15)^2 (.85)^4 = \boxed{0.176}$

لاحظ أن P أصبحت q

حل آخر نبقى على المتغير X ونقول احتمال 2 لم يصيبا يعني احتمال 4 أصابوا وتكون المطلوب

$f(X=4) = {}^{6}C_{4} (.85)^4 (.15)^2 = \boxed{0.176}$

أي نفس الناتج

Poisson Dist توزيع بواسون

وهو من التوزيعات المنفصلة التي لها تطبيقات مهمة، حيث يمثل م.ع X في هذه الحالة عدد مرات حدوث حدث معين في فترة زمنية معينة أو مساحة معينة، ومن أمثلته عدد حوادث السير في مدينة ما في وقت معين أو عدد المكالمات التي يستقبلها مكتب خلال ساعات الدوام، أو عدد الأخطاء المطبعية في كتاب معين... الخ. ففي مثل هذه الحالات يأخذ م.ع X القيم ∞ ,... ,2 ,1 ,0 ويعتمد في توزيعه على معلمة واحدة هي معدل الحدوث في الفترة الزمنية أو في المساحة المعينة ويرمز لهذا المعدل بالرمز (λ)

ونقول $X \sim P_o(\lambda)$

ودالته الاحتمالية هي:

$$\boxed{f(x) = p(X=x) = \frac{e^{-\lambda} \lambda^{x}}{x!}} \quad x = 0,1,2,3,\ldots\infty$$

ويتميز هذا التوزيع بتساوي كلاً من المتوسط والتباين

أي أن $\boxed{\mu = \sigma^2 = \lambda}$

$= \frac{e^{-\lambda} \lambda^{x}}{x!}$ ... احتمال استقبال ... مكالمات ... ساعة

$X(X=3)$

$\lambda = 2$

$= \frac{e^{-2} 2^{3}}{}$ ... $f(x=3)$

مثال / إذا كانت العيوب في القماش في شركة الغزل والنسيج تحدث بمعدل 6 عيوب في كل 100 م², فأوجد :-

(i) احتمال أن عدد العيوب 2 فقط في 100م² من هذا القماش ؟

(ii) = = = يزيد عن 2 في 100م² من هذا القماش ؟

(iii) = = 3 عيوب في قطعة مكونة من 300 م² = = = ؟

(iv) متوسط وتباين عدد العيوب في كل 500 م² ؟

الحل

(i) $p(x=2) = \frac{e^{-6}\,6^2}{2!} = 18\,e^{-6} = \boxed{0.045}$

(ii) $p(x>2) = 1 - p(x \le 2) = 1 - [p(x=0) + p(x=1) + p(x=2)]$

$= 1 - \left[\frac{e^{-6}\,6^0}{0!} + \frac{e^{-6}\,6^1}{1!} + \frac{e^{-6}\,6^2}{2!}\right]$

$= 1 - [1+6+18]\,e^{-6} = 1 - 25e^{-6} = \boxed{0.938}$

(iii) في هذه الحالة يجب تعديل المعدل λ ليصبح

$100\,م^2 \;—\; 6$

$300\,م^2 \longrightarrow ?$ $\quad \therefore ? = \frac{6 \times 300}{100} = 18$ عيب

$\therefore \lambda = 18$ عيب $/ 300\,م^2$ وبذلك يكون

$p(x=3) = \frac{e^{-18}\,18^3}{3!} = \boxed{0.000014}$

(iv) في هذه الحالة المساحة 500 م² إذن

$\lambda = \;6 \;—\; 100\,م^2$

$? \longleftarrow 500\,م^2$

$\therefore ? = \frac{500 \times 6}{100} = 30$ عيب $/ 500\,م^2$

$= \lambda$ ، $\sigma^2 = \mu$

$\therefore$ المتوسط = التباين = 30 عيب في كل 500م²

2017.10.17

ح 54

التوزيع الطبيعي Normal Dist

ويعتبر من أهم التوزيعات الاحتمالية المتصلة وذلك لأنه:

1- كثيرًا من المتغيرات العشوائية الغير طبيعية يمكن تحويلها إلى توزيع طبيعي بشروط معينة.

2- كثيرًا من الظواهر تتوزع توزيعًا طبيعيًا.

3- يدخل في كافة المجالات الصناعية، والزراعية والطبية والاقتصادية... الخ

تعريف:- إذا كان X م.ع يتبع التوزيع الطبيعي، فإن دالة كثافة احتماله لها الصيغة التالية:-

$$f_X(x) = P(X=x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}(x-\mu)^2/\sigma^2} \qquad \boxed{-\infty < x < \infty}$$

حيث (μ, σ^2) هما معلمتي التوزيع وهما المتوسط والتباين

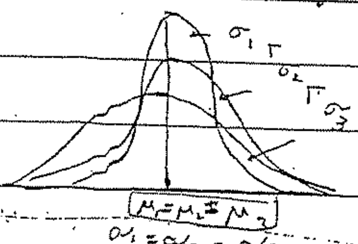
ونقول $X \sim N(\mu, \sigma^2)$، ومن أهم خصائصه:-

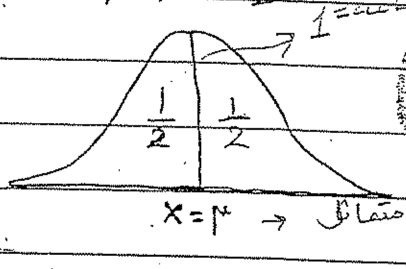
1) أن المساحة تحت منحنى التوزيع الطبيعي تساوي 1 أي أنه:-

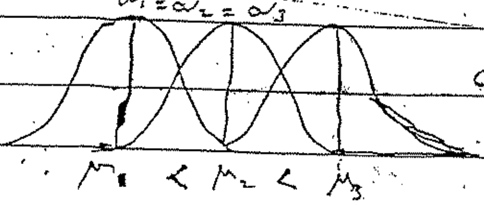
$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}(x-\mu)^2/\sigma^2} = 1 \qquad -\infty < x < \infty$$

2) أنه متماثل حول المتوسط μ أي أن:-

$$P(X \le \mu) = P(X \ge \mu) = \frac{1}{2}$$







3- يتحدد شكله بمعلميه μ & σ^2

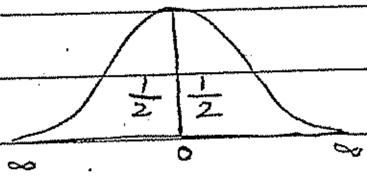
4- القيمة المتوقعة $E(X)=\mu$ و التباين $\sigma^2 = V(X)$.

نظرية :- مهمة جداً إذا كان X م.ع يتبع التوزيع الطبيعي بمتوسط μ وتباين σ^2 أي أن $X \sim N(\mu, \sigma^2)$

فإن المتغير العشوائي $\boxed{Z = \frac{X-\mu}{\sigma}}$ يتبع توزيع طبيعي بمتوسط صفر وتباين 1 أي أن :-

if $X \sim N(\mu, \sigma^2) \Rightarrow Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

ويقال بأنه م.ع Z يتبع التوزيع الطبيعي المعياري «قياسي»



الإشارة (Z) أول نوع

(البرهان 1-0)

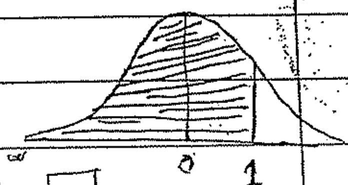
وتستخدم في إيجاد الاحتمالات المتعلقة بالتوزيعات الطبيعية الأخرى وتوجد جداول لذلك.

مثال إذا كان $X \sim N(80, 25)$ أوجد $P(X \le 85)$ ، $P(X \ge 80)$ ، $P(78 < X < 82)$

الحل

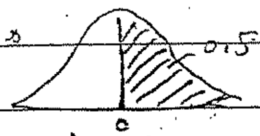
$P(X < 85) = P\left(Z < \frac{85-80}{5}\right) = P(Z \le 1)$

$= 0.8413$



$P(X \ge 80) = P\left(Z \ge \frac{80-80}{5}\right) = P(Z \ge 0) = \boxed{0.5}$

«من خاصية التماثل»

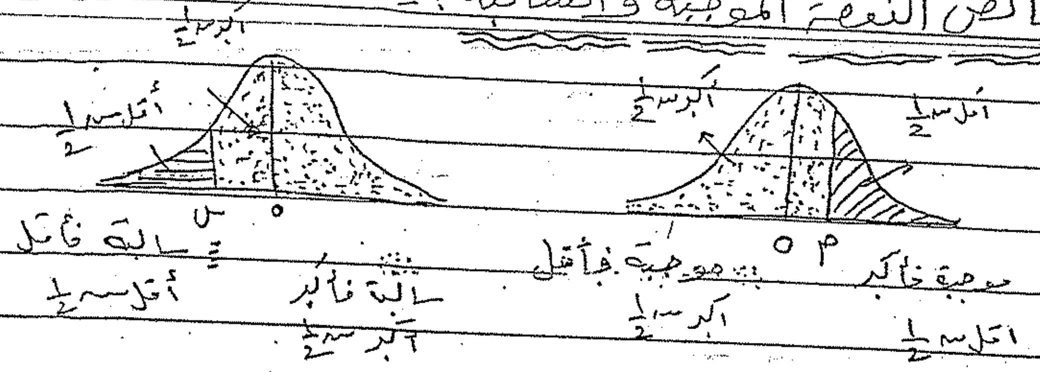


$P(78 \le X < 82) = P\left(\frac{78-80}{5} < Z \le \frac{82-80}{5}\right)$

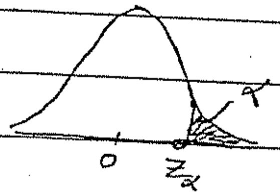
$= P(0.4 \le Z \le 0.4)$

$= \boxed{0.3108}$

خصائص النقطة الموجبة والسالبة :-



مثال: إذا كان Z_α ترمز إلى قيمة م.ع Z التي تقع على يمينها مساحة تساوي "α" أي أن :- $P(Z > Z_\alpha) = \alpha$



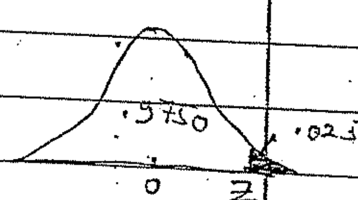
أوجد $Z_{.025}$ & $Z_{.9115}$

الحل

$Z_{.025} = P(Z > Z_1) = .025$ (أكبر — أقل من 1/2)

∴ Z_1 موجبة

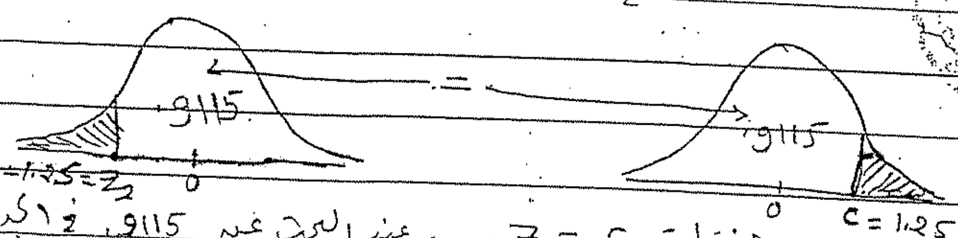
$\Rightarrow P(Z < Z_1) = .9750$



من الجدول نجد أن $Z_1 = 1.96$

$P(Z > Z_2) = .9115$ (أكبر — أكبر من 1/2)

∴ Z_2 نقطة سالبة كما في الشكل



حيث النقطة $c = -Z_2$ وعند البحث عن .9115 في الجدول نجد أن

$\boxed{c = 1.25}$

تابع 57

مثال إذا كانت أوزان مادة كيميائية تعبأ في زجاجات تتبع التوزيع الطبيعي بمتوسط 4.8 جرام وانحراف معياري 0.5 جرام. فما هي نسبة المئوية للزجاجات التي تحتوي على:
1- أقل من 4 جرامات ، (2) أكبر من 5 جرامات
3- ما بين 3.2 و 5.2 جرام من المادة الكيميائية ؟

الحل

نفرض أن X م.ع تمثل أوزان الزجاجات ∴

$$X \sim N(4.8, (.5)^2)$$

3

$$\therefore P(3.2 < X < 5.2) = P\left(\frac{3.2-4.8}{0.5} < Z < \frac{5.2-4.8}{0.5}\right)$$

$$P(-3.2 < Z < 0.8) = 0.4993 + 0.2881$$

$$= \boxed{0.7874}$$

يترك ① و ② للطالب.

لنفرض أنه طلب منا معرفة ما هو أقل وزن من المادة الكيميائية التي كانت نسبة الزجاجات التي تحتويه هي 2.5% ؟

الحل

نفرض أن هذا الوزن هو K ∴ من المعطيات يكون المطلوب

$$P(X > K) = 2.5\%$$

$$P\left(Z > \frac{K-4.8}{0.5}\right) = .025$$

نفرض أن $Z_1 = \frac{K-4.8}{0.5}$

$$\therefore P(Z > Z_1) = .025$$

من الجدول تكون $Z_1 = 1.96$

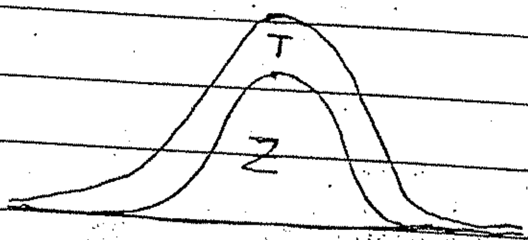
بالرجوع إلى قيمة Z_1 تكون

$$1.96 = \frac{K-4.8}{0.5}$$

$$\therefore K = 4.8 + 1.96(.5) = \boxed{5.78 \text{ جرام}}$$

توزيع T :

هو توزيع لمتغير عشوائي متصل يشبه التوزيع الطبيعي من حيث أنه متماثل حول الصفر ومنحنى التوزيع (T) يشبه منحنى (Z) إلا أن طرفيه يبتعدان قليلاً عن محور السينات كما في الشكل وتزداد دقة التقريب كلما ازداد حجم العينة ، ودالة كثافة توزيع (T) لها معلمة واحدة وهي (ν) وتمثل درجة الحرية حيث $\nu = n-1$



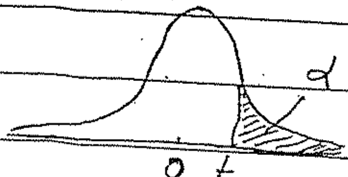
ودالة كثافة (T) معطاة بالمعادلة

$$f(t) = c\left(1+\frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

$-\infty < t < \infty$

حيث c عدد ثابت يعتمد على ν ليجعل المساحة تحت المنحنى تساوي 1 وكلما زادت ν اقترب T من Z

كيفية استخدام جدول T .

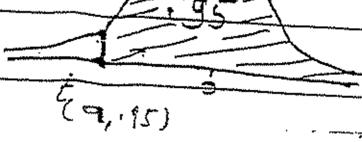


مثال $t_{(9,.05)} = 1.833$ مباشرة من الجدول

أما $t_{(9,.95)}$ والتي تعني أن $P(T > t_{(9)}) = .95$ أكبر

وهنا نعرف أن النقطة سالبة لأنها على يسار الصفر

والتي تساوي $t_{(9,.05)}$ - حيث لابد أن تكون سالبة $\boxed{-1.833}$



مثال أوجد $t_{(14,.99)}$

الحل

$$t_{(14,.99)} = -t_{(14,.01)} = -2.624$$

$$t_{(11,.005)} = 3.106$$

$$t_{(15,.975)} = -t_{(15,.025)} = -2.131$$

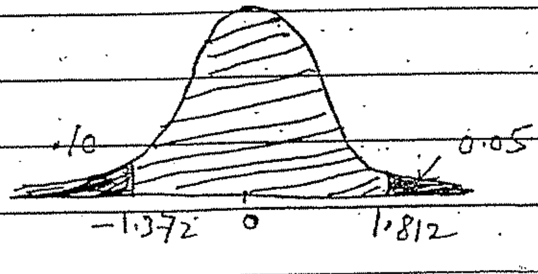
والقاعدة العامة هي

$$\boxed{t_{(\nu,1-\alpha)} = -t_{(\nu,\alpha)}}$$

تمرين: أوجد $P(-1.372<T<1.812)$ اذا كانت $\nu=10$

الحل

بالاستعانة بالرسم نجد أن المطلوب المساحة المظللة



والتي تساوي المساحة الكلية - (مجموع المساحة على يمين 1.812 وعلى يسار -1.372)

من الجدول T عند $\nu=10$ نجد أن

$$t_{(10,.05)} = 1.812$$

∴ المساحة على يمين 1.812 هي 0.05

والمساحة على يسار -1.372 = المساحة على يمين 1.372 [أقامتها]

ومن جدول T عند $\nu=10$ نجد أن $t_{(.10,10)} = 1.372$

$\therefore$ المساحة على يسار $-1.372 = .10$

$\therefore$ المساحة المظللة المطلوبة هي $1-(.05+.10)$

$= 1-.15 = .85$

$\therefore P(-1.372 < T < 1.812) = \boxed{.85}$

تمارين

1- أوجد قيمة $P(1.330 \le T \le 2.552)$ عندما تكون $\nu=18$

2- = = k بحيث $P(|T| \le k) = .90$ = = $\nu=14$

3- إذا كان $X \sim N(50, 9)$ أوجد قيمة الثابت K بحيث

$$P(X<K) = .95$$

4- إذا كانت السرعة في الطريق السريع على شكل التوزيع الطبيعي بمتوسط 100 كم/الساعة وبانحراف معياري 10 كم وكانت نسبة السائقين الذين قد خالفوا لتجاوزهم السرعة القصوى هي 9.34% فما هي السرعة القصوى المحددة لهذا الطريق؟

5- إذا كانت أطوال قطع خشب مصنع تتبع التوزيع الطبيعي بمتوسط μ وبانحراف معياري 3 أوجد قيمة هذا المتوسط إذا علمت أن احتمال أن ينتج المصنع قطعة طولها يقل عن 150 سم هو 2.5%؟

توزيعات المعاينة

تعريف المعاينة: هي اسلوب إحصائي يتم بواسطته الحصول على معلومات عن المجتمع ومعالمه (المتوسط، التباين، النسبة، ...)، من خلال دراسة جزء من المجتمع يسمى العينة، ويجب أن تكون العينة ممثلة للمجتمع الذي سحبت منه أي أن لها خصائص المجتمع الذي سحبت منه حتى تعطي معلومات صحيحة عن المجتمع، ومن أهم طرق المعاينة الاحتمالية ما يعرف بالعينة العشوائية البسيطة والتي يكون احتمال اختيار أي مفردة (مشاهدة) من مفردات المجتمع متساوي لكل المفردات، وأي قيمة يتم حسابها من مفردات العينة تسمى (إحصاءة).

نظرية: إذا أخذت عينة عشوائية بحجم (n) من مجتمع متوسطه μ وتباينه σ^2 فإن $\bar{x}$ يعتبر مقدر غير متحيز لمتوسط المجتمع μ وبتباين قدره $\frac{\sigma^2}{n}$، وكذلك تباين العينة "S^2" مقدر غير متحيز لتباين المجتمع أي أن:

$$E(\bar{x}) = \mu \quad \& \quad \sigma^2_{\bar{x}} = \sigma^2/n \quad \& \quad E(S^2) = \sigma^2$$

وتزداد جودة التقدير بزيادة حجم العينة

مثال بسيط لتوضيح ذلك: إذا كان لدينا مجتمع حجمه 5 وهو يمثل أعمار خمسة أطفال $x_1 = 6$ ، $x_2 = 8$ ، $x_3 = 10$ ، $x_4 = 12$ ، $x_5 = 14$

فإن جميع العينات الممكنة أخذها من هذا المجتمع بحجم 2، والمتوسطات المقابلة لها عليه وضعت في الجدول التالي وكما هو معلوم أن عدد الطرق هو $C^5_2 = 10$ مبينة كما يلي:

العينة	(6,8)	(6,10)	(6,12)	(6,14)	(8,10)	(8,12)	(8,14)	(10,12)	(10,14)	(12,14)
المتوسط الحسابي لها $\bar{x}$	$\bar{x}_1 = 7$	$\bar{x}_2 = 8$	$\bar{x}_3 = 9$	$\bar{x}_4 = 10$	$\bar{x}_5 = 9$	$\bar{x}_6 = 10$	$\bar{x}_7 = 11$	$\bar{x}_8 = 11$	$\bar{x}_9 = 12$	$\bar{x}_{10} = 13$

$$\mu = \frac{6+8+10+12+14}{5} = \frac{50}{5} = \boxed{10}$$

نفس النتائج

وكذلك $\dfrac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \cdots + \bar{x}_{10}}{10} = \dfrac{100}{10} = \boxed{10}$

نظرية ١ :- إذا كانت X م ع له توزيع طبيعي بمتوسط μ وتباين σ^2 فإن $\bar{X}$ يتوزع توزيعاً طبيعياً بمتوسط μ وتباين قدره σ^2/n أي أن :-

$$\text{if } X \sim N(\mu, \sigma^2) \Rightarrow \bar{X} \sim N(\mu, \sigma^2/n)$$

وعليه فإن ← (٢)

$$\Rightarrow \boxed{Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}} \sim N(0,1)$$

نظرية النهاية المركزية :- Central limit theorm

إذا أخذت عينة عشوائية بحجم n من مجتمع يتوزع (أي توزيع) بمتوسط μ وتباين σ^2 فإن متوسط العينة $\bar{X}$ يتوزع تقريباً توزيعاً طبيعياً بمتوسط μ وتباين $\frac{\sigma^2}{n}$ أي أن

$$\bar{X} \simeq N(\mu, \sigma^2/n)$$

الإشارة $\simeq$ تعني أن التوزيع تقريبي وعليه فإن :-

$$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \simeq N(0,1)$$

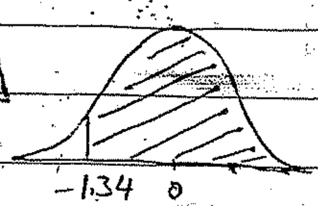
يشترط أن يكون حجم العينة كبير بدرجة كافية وعدم تحديد حجم العينة بسبب أن الحجم يعتمد على نوعية التوزيع الأصلي وعلى مدى ابتعاده عن التوزيع الطبيعي ولكن في الإجمال من خلال التطبيقات يكون من المعقول أن يكون $n > 30$ كافية لاستخدام نظرية النهاية المركزية.

مثال ١ :- إذا كانت X م ع تمثل درجات الذكاء لطلبة كلية ما يتبع $N(115, 100)$ ، فإذا شارك في اختبار الذكاء عشرون طالبا اختيروا بطريقة عشوائية ، أوجد احتمال أن يزيد متوسط درجات ذكائهم عن 112 درجة ؟ الحل

المطلوب هو $P(\bar{X} > 112)$

$$\therefore P\left(Z > \frac{112-115}{10/\sqrt{20}}\right) = P(Z > -1.34)$$

$$= 0.5 + 0.4099 = \boxed{0.9099}$$



مثال / إذا كان متوسط وتباين قوة قضبان الصلب المستخدمة من قبل شركة إنشاءات هندسية هما 75 كجم وتباين 160 كجم²، أوجد احتمال أن يتراوح متوسط قوة عينة عشوائية بحجم 64 من هذه القضبان بين 70 و 80 كجم ؟

الحل

حيث $n = 50 > 30$ سوف نطبق نظرية النهاية المركزية

فإذا كان X تمثل قوة القضبان فإن

$$\bar{X} \simeq N(\mu, \sigma^2/n)$$

ويكون المطلوب $P(70 < \bar{X} < 80)$

$$= P\left(\frac{70-75}{12.65/\sqrt{64}} < Z < \frac{80-75}{12.65/\sqrt{64}}\right)$$

$$= P\left(\frac{-5 \times 8}{12.65} < Z < \frac{5 \times 8}{12.65}\right) = P(-3.16 < Z < 3.16)$$

$$= 2 \times .4992 = \boxed{0.9984}$$

مثال / إذا كان متوسط مجتمع معين هو 99 وانحراف معياري 7

أوجد $E(\bar{X})$ & $V(\bar{X})$ لعينة بحجم $n=4$ & $n=25$

الحل

(1) $n=4$ $\quad E(\bar{X}) = \mu = 99$

$$V(\bar{X}) = \sigma^2/n = \frac{49}{4} = \boxed{12.25}$$

(2) $n=25$ $\quad E(\bar{X}) = 99$ & $V(\bar{X}) = \frac{49}{25} = \boxed{1.96}$

وإذا كان تباين المجتمع σ^2 مجهولا وحجم العينة (n) صغيراً، فإن الإحصائة

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

تتبع توزيع t بدرجة حرية $\nu = n-1$ (الذي سبق التعريف به انظر ص 58).

وعليه عكس أن نلخص توزيع متوسط العينة ($\bar{X}$) ثلاث حالات:

(1) σ^2 (معلومة) $\Rightarrow \bar{X} \sim N(\mu, \sigma^2/n) \Rightarrow Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$.

(2) σ^2 (مجهولة) + $n \geq 30 \Rightarrow \bar{X} \sim N(\mu, S^2/n) \Rightarrow Z = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim N(0,1)$

(3) σ^2 (مجهولة) + $n < 30 \Rightarrow \bar{X} \sim T = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{(n-1)}$.

توزيع الفرق ما بين متوسطي عينتين $(\bar{X}_1 - \bar{X}_2)$

في كثير من الأحيان عند دراسة مجتمعين يكون التركيز على معرفة فيما إذا كان هناك فرق ما بين متوسطي مجتمعين أو معرفة مقدار هذا الفرق فمثلا وجود برنامجين مختلفين للتدريب على وظيفة وترغب الجهة المختصة في معرفة أي البرنامجين في المتوسط يؤهل عاملين أكثر كفاءة، وللإجابة على ذلك جرت العادة باختيار عينتين (أي عينة من كل مجتمع) ثم مقارنة الفرق.

نظرية: إذا كانت $x_{11}, x_{12}, x_{13}, \ldots, x_{1n_1}$ تمثل عينة من مجتمع إحصائي يخضع لتوزيع متوسطه μ_1 وتباينه σ_1^2 وكانت $x_{21}, x_{22}, x_{23}, \ldots, x_{2n_2}$ عينة من مجتمع آخر متوسطه μ_2 وتباينه σ_2^2، وكان حجم العينتين كبيراً، فإن توزيع المعاينة للفرق ما بين المتوسطين $(\bar{X}_1 - \bar{X}_2)$ يتوزع وفق التوزيع الطبيعي بمتوسط قدره $\mu_1 - \mu_2$ وتباين قدره $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$ أي أن:

$$(\bar{x}_1-\bar{x}_2) \sim N\left(\mu_1-\mu_2 \;,\; \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)$$

ومنها

$$\boxed{Z = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim N(0,1)} \rightarrow (2)$$

مثال/ أخذت عينة عشوائية بحجم 20 من قيم متغير عشوائي X يخضع لتوزيع طبيعي بمتوسط 5 وتباين 0.5 ، وعينة أخرى مستقلة عن العينة الأولى من قيم متغير عشوائي Y يخضع لتوزيع طبيعي بمتوسط 4 وتباين 0.25 بحجم n_0 أيضا . أوجد :-

(i) $P(y-3>0)$ (ii) $P(\bar{y}>3)$ (iii) $P(\bar{x}-\bar{y}\le 0.5)$

الحل

$X \sim N(5 \;\; 0.5)$ & $n_1=20$

$Y \sim N(4 \;\; 0.25)$ & $n_2=20$

$\therefore$ (i) $P(y-3>0) = P(y>3) = P\left(Z>\frac{3-4}{0.5}\right)$

$= P(Z>-2) = P(Z<2)$

$= 0.5+0.4772 = \boxed{0.9772}$

(ii) $P(\bar{y}>3) = P\left(Z>\frac{3-4}{.5/\sqrt{20}}\right) = P(Z>-2\sqrt{20}) = 1$

(iii) $P(\bar{x}-\bar{y}\le 0.5)$

$= P\left(Z \le \frac{0.5-(5-4)}{\sqrt{\frac{.5}{20}+\frac{.25}{20}}}\right)$

$= P(Z\le -2.58) = P(Z\ge 2.58)$

$= 0.5-0.4951 = \boxed{0.0049}$

مثال: إذا علمت أن الإنتاج السنوي لأحد مناجم الذهب يتبع توزيع طبيعي بمتوسط 150 طن وانحراف معياري يساوي 20 طن بينما الإنتاج السنوي لمنجم آخر يتبع التوزيع الطبيعي بمتوسط 125 طن وانحراف معياري 25 طن، فإذا تم اختيار عينة من انتاج 5 أشهر من كل منجم بما يلي:

1) أن يكون متوسط عينة المنجم الأول أصغر أو يساوي متوسط عينة المنجم الثاني؟

2) أن يكون الفرق ما بين متوسطي العينتين أكبر من أو يساوي 60 طن؟

الحل:

نفرض أن $\bar{X}_1$ تمثل متوسط عينة المنجم الأول و $n_1=5$

$\bar{X}_2$ = = = = الثاني و $n_2=5$

لأن $\bar{X}_1-\bar{X}_2 \sim N\left(\mu_1-\mu_2, \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)$

$$\mu_{\bar{X}_1-\bar{X}_2} = \mu_1-\mu_2 = 150-125 = 25$$

$$\sigma^2_{\bar{X}_1-\bar{X}_2} = \frac{(20)^2}{5}+\frac{(25)^2}{5} = 205 \therefore \sigma_{\bar{X}-\bar{X}} = 14.32$$

now

$$P(\bar{X}_1 \le \bar{X}_2) = P(\bar{X}_1-\bar{X}_2 \le 0) \therefore P\left(Z \le \frac{0-(25)}{14.32}\right)$$

$$= P(Z \le -1.75) = P(Z \ge 1.75)$$

$$= 0.5-.4599 = \boxed{.0401}$$

أي أن فرصة أن يكون متوسط انتاج المنجم الأول أقل من أو يساوي متوسط انتاج المنجم الثاني تقريباً 4%

2) $$P(\bar{X}_1-\bar{X}_2 \ge 60) = P\left(Z \ge \frac{60-25}{14.32}\right) = P(Z \ge 2.44)$$

$$= 0.5-0.4927 = \boxed{.0073} = \boxed{0.01}$$

أي أن فرصة أن يكون الفرق بين متوسطي العينتين أكبر من أو يساوي 60 طن يساوي 1% تقريباً.

توزيع الفرق بين متوسطات عينتين في حالة σ_1^2 و σ_2^2 مجهولتان

وحجم العينتين صغير فإن توزيع المعاينة للفرق بين

متوسطي عينتين يتوزع وفق توزيع t بدرجة حرية n_1+n_2-2

حيث

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

حيث

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

ويسمى التباين المشترك

مثال/ نوعان من محطات التوليد للطاقة، حيث كانت تكلفة الصيانة السنوية للنوع الأول تتبع $N(30, \sigma^2)$ والثانية $N(35, \sigma^2)$ اخذت عينة من النوع الأول حجمها 8 محطات فكان متوسط تكلفة صيانتها بالآلاف الدينارات 27 د، وانحراف معياري 40 د، وكذلك عينة من النوع الثاني بحجم 15 محطة فكان متوسط تكلفة الصيانة لـ 35 د بانحراف معياري 18 د، أوجد احتمال أن الفرق بين متوسطي التكلفة بين العينتين لا يزيد عن 10 آلاف دينار؟

الحل

$n_1=8 \quad \bar{X}_1=27 \quad S_1=40$

$n_2=15 \quad \bar{X}_2=35 \quad S_2=18$

$\because$ n_1 و n_2 صغيرتان $(n_1+n_2-2<30)$ و $\sigma_1^2=\sigma_2^2=\sigma^2$ مجهولتان ومتساويتان نستخدم الاحصاءة T السابقة

نوجد اولاً S_p حيث

$$S_p = \sqrt{\frac{7(40)^2+14(18)^2}{8+15-2}} = \boxed{27.37}$$

$$\therefore P(\bar{X}_1-\bar{X}_2<10) = P\left(T<\frac{10-(30-35)}{27.37\sqrt{\frac{1}{8}+\frac{1}{15}}}\right) = P(T<1.25)$$

$$= P(T<1.3) \simeq 0.90$$

أي أن احتمال أن الفرق في التكاليف بين متوسطي القسمين لا يزيد عن 15 ألف دينار هو تقريباً 95 %.

توزيع المعاينة لنسبة العينة $\hat{p}$

إذا كانت X تمثل عدد عناصر مجتمع إحصائي تمتلك صفة معينة وكانت N تمثل عدد عناصر ذلك المجتمع فإن نسبة هذه العناصر التي تمتلك تلك الصفة ويرمز لها بالرمز P حيث $P=\frac{X}{N}$

وإذا كانت $x_1, x_2, x_3, \ldots\ldots x_n$ تمثل عينة من هذا المجتمع وكانت x تمثل عدد العناصر في هذه العينة التي تمتلك تلك الصفة وكان حجم العينة n، فإن نسبة هذه العناصر في العينة ويرمز لها بالرمز $\hat{p}$ حيث $\hat{p}=\frac{x}{n}$ ويكون توزيع $\hat{p}$ تقريباً وفق التوزيع الطبيعي بشرط أن n كبيرة & $np \geq 5$ وبالتالي فإن

$$\boxed{Z=\frac{\hat{p}-P}{\sqrt{\frac{P(1-P)}{n}}} \sim N(0,1)}$$

لاحظ أن

$E(\hat{p})=\mu_{\hat{p}}=P$

$\sigma^2_{\hat{p}}=\frac{P(1-P)}{n}$

مثال: إذا علمت أن 35% من الحوادث كانت بسبب السرعة الفائقة، فإذا تم اختيار عينة عشوائية من 100 حادث مسجلة بدوائر المرور أوجد:

1) احتمال أن 45% فأكثر كانت بسبب السرعة الفائقة؟

2) من 35% إلى 50% من الحوادث كانت بسبب السرعة الفائقة؟

الحل

$P=0.35$ النسبة العامة في المجتمع، & $n=100$ حجم العينة

ويكون المطلوب

① $P(\hat{P} \geq 0.45)$

$$= P\left(Z \geq \frac{0.45-0.35}{\sqrt{\frac{0.35 \times .65}{100}}}\right) = P(Z \geq 2.1)$$

$$= 0.5 - 0.4821 = \boxed{0.0179}$$

② $P(0.35 \leq \hat{P} \leq 0.50)$

$$= P\left(\frac{0.35-0.35}{\sqrt{\frac{0.35 \times .65}{100}}} \leq Z \leq \frac{0.5-0.35}{\sqrt{\frac{.35 \times .65}{100}}}\right)$$

$$= P(0 \leq Z \leq 3.14) = \boxed{0.4992}$$

مثال / إذا علمت أن نسبة العمال الذين لهم حساسية من نوع معين من الروائح 5٪ ، فإذا تم اختيار عينة حجمها 150 من العمال فما احتمال:

1) أن تكون نسبة العمال في هذه العينة الذين لهم حساسية من هذا النوع من الروائح أكبر من 3٪ ؟

2) أن تكون نسبة العمال الذين لهم حساسية من هذا النوع من الروائح تقل عن 6٪

الحل

(1) $P(\hat{P} > 0.03) = P\left(Z > \frac{0.03-0.05}{\sqrt{\frac{.05 \times .95}{150}}}\right)$

$$= P(Z > -1.12) = 0.5 + 0.3686 = \boxed{0.8686}$$

② $P(\hat{P} < .06) = P\left(Z < \frac{.06-0.05}{\sqrt{\frac{.05 \times .95}{150}}}\right)$

$$= P(Z < 0.56) = 0.5 + 0.2123 = \boxed{0.7123}$$

تمرين: في المثال السابق أوجد احتمال أن نسبة العمال الذين ليس لهم حساب من هذا المصنع في البنوك لا تزيد عن 90%؟

(لا يترك الحل للطالب).

توزيع المعاينة للفرق بين نسبتي عينتين $\hat{p}_1 , \hat{p}_2$

إذا كانت $x_1, x_2, \ldots\ldots, x_{n_1}$ تمثل عناصر عينة من مجتمع نسبة عناصره التي لها خاصية مدار البحث P_1 & ونسبة العناصر في العينة التي لها نفس الخاصية هي $\hat{p}_1$، وكانت $y_1, y_2, \ldots\ldots, y_{n_2}$ هي عناصر عينة من مجتمع آخر له نسبة عناصر لنفس الخاصية هي P_2 ونسبة العناصر التي لها نفس الخاصية في العينة هي $\hat{p}_2$ وكانت العينتين مستقلتين وحجمهما كبير فإن توزيع المعاينة للفرق بين نسبتي العينتين $\hat{p}_1 - \hat{p}_2$ يتبع تقريبا التوزيع الطبيعي بمتوسط $\mu_{\hat{p}_1-\hat{p}_2} = P_1 - P_2$ وتباين $\sigma^2_{\hat{p}_1-\hat{p}_2} = \frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}$

أي أن
$$\hat{p}_1 - \hat{p}_2 \simeq N\left(P_1 - P_2 , \frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}\right)$$

وعليه فإن

$$\boxed{Z = \frac{(\hat{p}_1 - \hat{p}_2) - (P_1 - P_2)}{\sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}} \simeq N(0,1)}$$

مثال: تستورد شركة الإلكترونيات قطع الغيار اللازمة لها من مصدرين مختلفين، وأن قطع الغيار الواردة من المصدر الأول ترفض بمعدل 8% بسبب بعض العيوب، بينما نسبة الرفض من المصدر الثاني 5%. فإذا كانت الخطوط الإنتاجية في الشركة تتحدد 150 قطعة

من المصدر الأول و 300 قطعة من المصدر الثاني يومياً. فما هي نسبة الأيام التي سوف يكون فيها الفرق بين نسبتي الرفض من المصدرين أقل من أو يساوي 7% ؟

الحل

نفرض أن $\hat{p}_1$ نسبة الرفض بالعينة من المصدر الأول و

$\hat{p}_2$ = = = = الثاني)

بما أن $n=300$ & $n=150$ كبيرين ∴

$$\hat{p}_1-\hat{p}_2 \simeq N\left(\mu_{\hat{p}_1-\hat{p}_2}, \sigma^2_{\hat{p}_1-\hat{p}_2}\right)$$

حيث

$$\mu_{\hat{p}_1-\hat{p}_2} = P_1 - P_2 = .08 - 0.05 = \boxed{.03}$$

$$\sigma^2_{\hat{p}_1-\hat{p}_2} = \frac{(.08)(0.92)}{150} + \frac{(.05)(.95)}{300} = .00065$$

ويكون المطلوب

$$P(\hat{p}_1-\hat{p}_2 \leq .07)$$

$$\therefore P\left(Z \leq \frac{.07-0.03}{\sqrt{.00065}}\right) = P(Z \leq 1.57) = 0.5 + .4418 = \boxed{0.9418}$$

أي أن حوالي 94% من الأيام سيكون فيها الفرق بين نسبتي الرفض من المصدرين أقل من أو يساوي 7% .

مثال: إذا كان 5% من إنتاج الخط الأول بمصنع للأدوية غير مطابق للمواصفات ($P_1=.05$) بينما نسبة الخط الثاني الغير مطابق للمواصفات هي 6% ، أخذت عينة عشوائية حجمها 400 قطعة من كل خط أوجد

(i) $P(|\hat{p}_1-\hat{p}_2| \leq .04)$ (ii) $P(|\hat{p}_1-\hat{p}_2| > .04)$

الحل

نفرض أن $\hat{p}_1$ تمثل نسبة الإنتاج الغير مطابق في العينة من الخط الأول
$\hat{p}_2$ = = = = = الخط الثاني
حيث $\hat{p}$ تمثل النسبة = = = =
لأن ن كبيرة (تقريبياً)

$$\mu_{\hat{p}_1-\hat{p}_2} = 0.05-0.06 = \boxed{-0.01}$$

$$\sigma^2_{\hat{p}_1-\hat{p}_2} = \frac{.05 \times .95}{400} + \frac{.06 \times .94}{400} = \boxed{0.00026}$$

$$\therefore \sigma_{\hat{p}_1-\hat{p}_2} = \boxed{.016}$$

(i) $P(|\hat{p}_1-\hat{p}_2| \le .04) \Rightarrow P(-.04 \le \hat{p}_1-\hat{p}_2 \le .04)$

$$\therefore P\left(\frac{-.04-(-.01)}{.016} \le Z \le \frac{.04-(-0.01)}{.016}\right)$$

$$= P(-1.88 \le Z \le 3.13) = .4699 + .4991$$

$$= \boxed{0.9690}$$

(ii) $P(|\hat{p}_1-\hat{p}_2| \ge .04) = 1 - P(|\hat{p}_1-\hat{p}_2| \le 0.04)$

$$= 1 - .9690 = \boxed{0.0310}$$

أي أن المطلوب في رقم (ii) هو مكملة المطلوب في رقم (i)

تعريف بالقيمة المطلقة مثلا $|x| < 2 \Rightarrow -2 < x < 2$

$|x| \ge 2 \Rightarrow (x \le -2 + x \ge 2)$

ج. التقدير C. Estimation

مما سبق بينا أن متوسط العينة $\bar{x}$ يعتبر مقدراً غير متحيز لمتوسط المجتمع μ بتباين قدره σ^2/n وكذلك تباين العينة S^2 مقدر غير متحيز لتباين المجتمع σ^2

أي أن
$$E(\bar{X})=\mu \quad \& \quad E(S^2)=\sigma^2$$

وتزداد جودة التقدير بازدياد حجم العينة ويسمى تقدير بالنقطة أي يتم تقدير معلمة مجتمع مجهولة بالإحصاءة المقابلة لها في العينة، وتكون الإحصاءة مقدراً جيداً إذا توفر فيها الشرطان

1) أن القيمة المتوقعة للإحصاءة مساوية لمعلمة المجتمع

أي أن
$$E(q) = Q$$

وأن يكون تباينها أصغر مقارنة بمقدرات أخرى

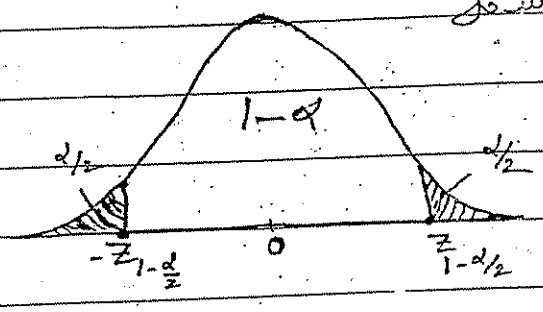
أي أن لأي مقدر آخر
$$\sigma_q^2 = E(q-Q)^2 < \sigma^2$$

وقد تقدر معلمة المجتمع المجهولة بفترة لها بداية ونهاية محسوبتان بالاعتماد على إحصاءات العينة، كأن تقدر المعلمة Q بأن نقول أنها تقع في الفترة (q_1, q_2) وتتميز الفترة بأن تكون مصاحبة بدرجة ثقة تبين درجة ثقتنا في أن هذه الفترة سوف تحتوي على المعلمة المجهولة، والقيمتان q_1, q_2 يطلق عليهما الحد الأدنى والحد الأعلى لفترة الثقة على التوالي، ويعرف معامل الثقة بأنه احتمال أن تحتوي فترة الثقة (q_1, q_2) على المعلمة المجهولة Q ويرمز له بالرمز $(1-\alpha)$ أي أن $P(q_1 < Q < q_2) = 1-\alpha$ وسنبدأ

بإيجاد أو تقدير فترة ثقة لمتوسط المجتمع μ كما يلي:

١- تقدير الفترة لمتوسط المجتمع μ بمعلومية σ^2

إذا كان $X \sim N(\mu, \sigma^2)$ فنحن نعلم من جداول التوزيع الطبيعي المعياري كما في الشكل



أي أن $P(-Z_{1-\frac{\alpha}{2}} < Z < Z_{1-\frac{\alpha}{2}}) = 1-\alpha \rightarrow \boxed{I}$

ح- $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ «كما أفترضنا سابقا» نعوض بها في المعادلة $\boxed{I}$
ينتج أن

$$P\left(-Z_{1-\frac{\alpha}{2}} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < Z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

$$\therefore P\left(-Z_{1-\frac{\alpha}{2}}\sigma/\sqrt{n} < \bar{X}-\mu < Z_{1-\frac{\alpha}{2}}\sigma/\sqrt{n}\right) = 1-\alpha$$

ومنه نحصل على

$$P\left(\bar{X}-Z_{1-\frac{\alpha}{2}}\sigma/\sqrt{n} < \mu < \bar{X}+Z_{1-\frac{\alpha}{2}}\sigma/\sqrt{n}\right) = 1-\alpha$$

وهذه الفترة $(\bar{x}-Z_{1-\frac{\alpha}{2}}\sigma/\sqrt{n}\ ,\ \bar{x}+Z_{1-\frac{\alpha}{2}}\sigma/\sqrt{n})$ تسمى $(1-\alpha)100\%$ فترة ثقة لمتوسط المجتمع μ ويمكن كتابتها على الصورة الآتية:

$$\boxed{\bar{x} \pm Z_{1-\frac{\alpha}{2}}\sigma/\sqrt{n}}$$

ويسمى $(1-\alpha)$ معامل أو مستوى الثقة

فمثلا إذا كانت $\alpha = .05$ فإن 95% فترة ثقة لـ μ هي

$\bar{x} \pm Z_{.975}\sigma/\sqrt{n}$

وإذا كانت $\boxed{\sigma^2 \text{ مجهولة}}$ وحجم العينة (n) كبير $n \geq 30$

فإن فترة الثقة لمتوسط المجتمع μ هي نفسها السابقة مع استبدال σ بـ s الانحراف المعياري للعينة
أي أن $(1-\alpha)$ فترة ثقة للمتوسط μ في هذه الحالة هي:

$$\boxed{\bar{x} \pm Z_{1-\frac{\alpha}{2}}\, s/\sqrt{n}}$$

* أما إذا كانت σ^2 مجهولة و $(n)<30$ فإن فترة الثقة لمتوسط المجتمع μ في هذه الحالة يتم حسابها باستخدام الإحصاءة $T=\frac{\bar{x}-\mu}{s/\sqrt{n}}$ وتكون على النحو التالي:

$$\boxed{\bar{x} \pm t_{(n-1,\frac{\alpha}{2})}\, s/\sqrt{n}}$$

الخلاصة: فترة ثقة لمتوسط المجتمع μ لها ثلاث حالات وهي:

① $(\bar{x} \pm Z_{1-\alpha/2}\, \sigma/\sqrt{n}) \rightarrow$ σ^2 معلومة

② $(\bar{x} \pm Z_{1-\frac{\alpha}{2}}\, s/\sqrt{n}) \rightarrow$ $n \geq 30$ (كبيرة) + σ^2 مجهولة

③ $(\bar{x} \pm t_{(n-1,\alpha/2)}\, s/\sqrt{n}) \rightarrow$ $n<30$ + σ^2 مجهولة

مثال// أخذت عينة عشوائية بحجم 100 مفردة من مجتمع متوسطه μ غير معلوم وتباينه $169 = \sigma^2$، فكان متوسط هذه العينة $\bar{x}=135$. أوجد 99٪ فترة ثقة لمتوسط المجتمع μ؟
الحل:

$1-\alpha = .99 \Rightarrow \alpha = .01 \Rightarrow \frac{\alpha}{2} = .005 \quad \therefore 1-\frac{\alpha}{2} = .995$

$\therefore Z_{.995} = 2.575$

بما أن σ^2 معلوم نستخدم الإحصاءة Z وتكون فترة الثقة لـ μ هي:

تمرين 7

$$\bar{x} \pm Z_{.995}\, \sigma/\sqrt{n} = 135 \pm 2.575 \times 13/10$$

$$= 135 \pm 3.35 = (131.65\, ,\ 138.35)$$

وتفسير ذلك كالتالي: إذا تم سحب جميع العينات الممكنة ذات الحجم 100 من نفس المجتمع وحسبت 99% فترة ثقة للمتوسط μ فإننا نتوقع 99% من هذه الفترات سوف تحتوي على قيمة μ

مصنع لإنتاج المكرونة يعبئ المكرونة في أكياس ذات وزن 500 جم وللتأكد من أن عملية الوزن تجري طبقا للمواصفات أخذت عينة عشوائية من 8 أكياس وحسبت أوزانها فكانت كما يلي:

496 ، 498 ، 501 ، 500 ، 495 ، 499 ، 503 ، 497

أحسب 95% فترة ثقة لمتوسط أوزان أكياس المكرونة المعبأة بهذا المصنع؟

الحل

نحسب أولاً $\bar{x} = \sum_{i=1}^{8} x_i / n = \frac{3989}{8} = 498.625$

ولذلك نحسب S^2 حيث

$$S^2 = \frac{1}{n-1}\left[\sum x_i^2 - n\bar{x}^2\right]$$

$$= \frac{1}{7}\left[1989065 - 8 \times 248626.89\right] = \boxed{7.125}$$

$\left[1-\alpha = .95 \Rightarrow \alpha = .05 \Rightarrow \frac{\alpha}{2} = .025\right]$ σ^2 مجهولة و صغيرة $n = 8$

نستخدم إحصاءة T وتكون فترة الثقة على النحو التالي:

$$\bar{x} \pm \frac{s}{\sqrt{n}}\, t_{(n-1, \alpha/2)} = 498.625 \pm \frac{\sqrt{7.125}}{\sqrt{8}}\, t_{(7, .025)}$$

$$= 498.625 \pm 0.94\,(2.36463)$$

$$= 498.625 \pm 2.223 = (496.4\, ,\ 500.8$$

أي أن متوسط أوزان أكياس المكرونة في هذا المصنع سوف يكون في الفترة (496.4 , 500.85) جرام بدرجة ثقة 95%

$496.4 < \mu < 500.85$ الحد الأدنى الحد الأعلى

فترة ثقة للفرق بين متوسطين مجتمعين $\mu_1 - \mu_2$ [σ_1^2, σ_2^2 معلومتان]

بالرجوع إلى توزيع المعاينة للفرق بين متوسطي عينتين نجد أن

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

وعليه فإن

$$P\left(-Z_{1-\frac{\alpha}{2}} < \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < Z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

ومنه كما سبق نجد أن $(1-\alpha)100\%$ فترة ثقة للفرق بين $\mu_1 - \mu_2$ هي:

$$\boxed{(\bar{x}_1 - \bar{x}_2) \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

وعندما تكون σ_1^2, σ_2^2 مجهولتان ومتساويتان أي أن $\sigma_1^2 = \sigma_2^2 = \sigma^2$:

وفي هذه الحالة يتم تقدير σ^2 بإيجاد ما يعرف بالتباين المشترك بين s_1^2 و s_2^2 والذي يساوي:

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

وبالرجوع إلى توزيع المعاينة للفرق بين متوسطي عينتين $\bar{x}_1 - \bar{x}_2$ في هذه الحالة نجد أن

$$P\left(-t_{(n_1+n_2-2, \frac{\alpha}{2})} < \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} < t_{(n_1+n_2-2, \frac{\alpha}{2})}\right) = 1-$$

ومنها تكون $(1-\alpha)100\%$ فترة ثقة لـ $\mu_1 - \mu_2$ هي:

$$\boxed{(\bar{x}_1-\bar{x}_2) \pm t_{(n_1+n_2-2,\ \alpha/2)}\, S_p \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

ملاحظة: إذا كان $(n_1+n_2-2 \geq 30)$ فإن فترة الثقة السابقة تتحول إلى:

$$\boxed{(\bar{x}_1-\bar{x}_2) \pm Z_{1-\frac{\alpha}{2}}\, S_p \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

مثال/ تم تدريب 10 متدربين على تجميع جهاز حاسوب باستخدام طريقة تدريب عادية (1) ، ثم تدريب 12 متدرب آخر على تجميع نفس النوع من الحاسوب باستخدام طريقة مبتكرة للتدريب (2) ، وبعد نهاية فترة التدريب حسب متوسط الزمن الذي استغرقه المتدربون في تجميع الجهاز بالساعات فكان $\bar{x}_1=36$ ، $\bar{x}_2=29$. فإذا كان زمن التجميع لكل متدرب يتوزع التوزيع الطبيعي $N(\mu_2,20)$ لـ 2 و $N(\mu_1,25)$ ، أحسب 99% فترة ثقة للفرق بين متوسطي زمن التجميع؟

الحل/

$1-\alpha=.99 \Rightarrow \alpha=.01 \therefore \frac{\alpha}{2}=.005 \Rightarrow 1-\frac{\alpha}{2}=.995$

∵ معلومتان σ_1^2 و σ_2^2

$\therefore Z_{.995}=2.575$

∴ فترة الثقة لـ $\mu_1-\mu_2$ هي:

$$(\bar{x}_1-\bar{x}_2)\pm Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} = (36-29)\pm 2.575\sqrt{\frac{25}{10}+\frac{20}{12}}$$

$= (1.74, 12.26)$ ساعة

10) نوعان من محطات لتوليد الكهرباء ، تكلفة الصيانة للنوع الأول تتبع التوزيع الطبيعي $N(\mu_1,\sigma^2)$ والثاني $N(\mu_2,\sigma^2)$ ، أخذت عينة من النوع الأول من 8 محطات فكان متوسط صيانتها السنوية بالآلاف الدينارات 27 ، انحراف معياري 4.0 ، وكذلك عينة من النوع الثاني بحجم 15

... محلم نكار متوسط مصاريفها بالدولار = 35 د وانحراف معياري 18. أوجد فترة ثقة للفرق بين متوسطي مصاريف النوعين $(\alpha=0.05)$

الحل

لاحظ أن $\sigma_1^2=\sigma_2^2=\sigma^2$ مجهولتان ومتساويتان وعليه نحسب أولا

$$S_p=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}=\sqrt{\frac{7(40)^2+14(18)^2}{21}}=27.37$$

فترة ثقة 95% للفرق بين المتوسطين هي:

$$\alpha=0.05 \quad \therefore \quad \frac{\alpha}{2}=0.025$$

$$(\bar{X}_1-\bar{X}_2)\pm t_{(21,0.025)}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$=(27-35)\pm 2.07963(27.37)\sqrt{\frac{1}{8}+\frac{1}{15}}$$

$$=-8\pm 24.92 \rightarrow (-32.92,\ 16.92)$$

إيجاد فترة ثقة للنسبة في المجتمع P: إذا ... بالرجوع إلى التوزيع العيني وجدنا أن

$$\hat{P}\sim N\left(P,\frac{P(1-P)}{n}\right)$$

وحيث أن P هنا مجهولة فإنه يتم تقديرها بـ $\hat{P}$ نسبة العينة وحيث في أغلب التطبيقات العملية يكون حجم العينة المستخدمة للتقدير كبير وبالتالي يمكننا تقدير الانحراف المعياري

$\sqrt{\frac{P(1-P)}{n}}$ بالتعويض عن قيمة P بقيمة $\hat{P}$ دون أن يفقد تقدير غير متحيز لها وبالتالي يكون

$$Z=\frac{\hat{P}-P}{\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}}\sim N(0,1)$$

ولإيجاد فترة الثقة $(1-\alpha)100\%$ للنسبة $\hat{P}$ نحسب

$$P\left(-Z_{1-\frac{\alpha}{2}} < Z < Z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

$$P\left(-Z_{1-\frac{\alpha}{2}} < \frac{\hat{p}-P}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} < Z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

ومنها نحصل على فترة ثقة لـ P على النحو التالي:

$$\boxed{\hat{p} \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$$

مثال: في دراسة لمعرفة تفضيل السائقين للسيارات صغيرة الحجم تم استبيان 500 سائق تم اختيارهم بطريقة عشوائية من إحدى المدن فقال 362 منهم أنهم يفضلون السيارات صغيرة الحجم. أحسب 95% فترة ثقة لنسبة السائقين الذين يفضلون سيارات صغيرة الحجم.

الحل:

$1-\alpha = .95 \therefore \alpha = .05 \Rightarrow \frac{\alpha}{2} = .025 \Rightarrow 1-\frac{\alpha}{2} = .975$

$Z_{.975} = 1.96$ و $\hat{p} = \frac{362}{500} = 0.72$

وتكون فترة الثقة هي:

$$0.72 \pm 1.96 \sqrt{\frac{0.72 \times .28}{500}}$$

$$= 0.72 \pm .04 \Rightarrow (.68 \quad .76)$$

أي أن نسبة الذين يفضلون السيارات الصغيرة هي لا تقل عن 68% ولا تزيد عن 76% بثقة قدرها 95%.

ثانياً: فترة ثقة للفرق بين نسبتين لمجتمعين $(P_1 - P_2)$:

بفرض أن نسبة الوحدات التي لا تحمل صفة معينة في المجتمع الأول هي P_1 وأن $\hat{p}_1$ هي نسبة الوحدات في عينة حجمها n_1 من هذا المجتمع لها نفس الخاصية وكذلك P_2 هي نسبة الوحدات التي لا نفس الخاصية في المجتمع

وأن $\hat{p}_2$ هي نسبة المفردات في عينة حجمها n_2 من المجتمع الثاني، فإن
فترة الثقة للفرق بين نسبتي المجتمعين $(P_1 - P_2)$ هي:

$$\boxed{(\hat{p}_1 - \hat{p}_2) \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

مثال/ أخذت عينة عشوائية من طلبة كلية العلوم بعدد 50 طالب فوجد أن 17 طالب لديهم جهاز حاسوب بالمنزل، وعينة أخرى عشوائية بعدد 80 طالب من كلية الهندسة فوجد أن 44 طالب لديهم جهاز حاسوب بالمنزل. أوجد فترة ثقة للفرق بين نسبة الطلاب الذين لديهم حاسوب في المنزل بين الكليتين.

الحل

$\hat{p}_1 = \frac{17}{50} = 0.34$ & $\hat{p}_2 = \frac{44}{80} = 0.55$ $\quad Z_{.975}$

$\therefore$ 95% فترة ثقة للفرق بين النسبتين $(P_1 - P_2)$ هي:

$$(.34 - .55) \pm 1.96 \sqrt{\frac{(.34)(.66)}{50} + \frac{(.55)(.45)}{80}}$$

$$= -0.21 \pm 0.17 = (-0.38 \,,\, -0.04)$$

تمرين/ نوعان من الثلاجات A و B ضمانة كل منهما سنة. في عينة عشوائية من 50 ثلاجة من النوع A تم تعطل 8 ثلاجات قبل نهاية السنة الضمانية، بينما في عينة مكونة من 70 ثلاجة من النوع B تعطلت منها 12 ثلاجة قبل نهاية سنة الضمان. احسب 95% فترة ثقة للفرق بين نسبتي تعطل النوعين قبل نهاية سنة الضمان؟

اختبارات الفروض

الفرض الإحصائي : هو جملة غالبًا ما تتعلق بالقيمة الحقيقية لإحدى معالم المجتمع المجهولة (كالمتوسط، التباين، النسبة ... الخ). وهذه الجملة أو الفرض قد يكون صحيحًا أو خاطئًا، ولذا يجب علينا اختباره بالاعتماد على عينة عشوائية مسحوبة من هذا المجتمع قيد الدراسة وكمثال على ذلك قد يريد مصنع للأدوية أن يختبر فيما إذا كانت نسبة المعاب في الإنتاج لا تزيد عن ٣٪ مثلاً، أو أن يختبر أن متوسط ما يصدره من إنتاج معين يزيد عن كمية معينة في الشهر.

فرض العدم والفرض البديل : Null & Alternative Hyps

يتكون الفرض الإحصائي من فرض العدم والذي يرمز له عادة بالرمز H_0، وهو يمثل حالة الوضع الراهن الذي يريد الباحث أن يرفضه، وفي حالة رفضه لفرض العدم، فإنه يقبل الفرض البديل المناقض له والذي يرمز له بالرمز H_1 والذي يمثل الحالة التي يرغب الباحث إثباتها.

مثلاً: إذا كان الفرض الإحصائي يتعلق باختبار فيما إذا كانت النسبة P لا تزيد عن 3٪ فإن هذا الفرض يكتب كما يلي:

$$H_0: P = .03$$

$$H_1: P < .03$$

وعلى الباحث أن يقوم بعملية اختبار لهذا الفرض الإحصائي، وعليه أن يقرر قبول أحد هذين الفرضين ورفض الآخر.

أنواع الأخطاء : Types of Errors

يمكن تلخيص ذلك في الجدول الآتي:

33

القرار / الفرضية	قبول H_0	رفض H_0
H_0: صحيح	✓	X
H_0: خاطئ	X	✓

خطأ من النوع الأول ← Type I error

خطأ من النوع الثاني Type II error

خطأ من النوع الأول: وهو رفض فرض العدم H_0 علماً بأنه صحيح

مستوى المعنوية (α): وهو احتمال الوقوع في الخطأ الأول أي أنه

$$P(\text{رفض } H_0 / H_0 \text{ صحيح}) = \alpha$$

خطأ من النوع الثاني: وهو قبول H_0 علماً بأنه غير صحيح واحتمال الوقوع فيه يرمز له بالرمز β أي أنه

$$P(\text{قبول } H_0 / H_0 \text{ خاطئ}) = \beta$$

قوة الاختبار: Power of Test ويرمز له بالرمز $1-\beta$ وهو احتمال رفض H_0 علماً بأنه خطأ أي أنه

$$P(\text{رفض } H_0 / H_0 \text{ خطأ}) = 1-\beta$$

منطقة الرفض: وهي مجموعة قيم إحصاءة الاختبار التي تؤدي إلى رفض H_0 وبالتالي فإن جميع القيم التي تؤدي إلى قبول H_0 تسمى منطقة القبول.

اختبار ذو جانب أو جانبين "طرف أو طرفين":

الذي يحدد فيما إذا كان الاختبار ذو جانب واحد أو جانبين هو الفرض البديل. فإذا كان الفرض البديل يحتوي على إشارة "$>$" أو "$<$" فإن الاختبار يقال له اختبار ذو طرف أو جانب واحد، أما إذا كان الفرض البديل يحتوي على إشارة "$\neq$" فيقال اختبار ذو طرفين أو جانبين. مثلاً $H_0: \mu=50$ ضد $H_1: \mu<50$ هو اختبار ذو طرف بينما الفرض $H_0: \mu=50$ ضد $H_1: \mu\neq50$ هو اختبار ذو طرفين.

خطوات اختبار الفرض الإحصائي:

1) صياغة فرض العدم والفرض البديل H_0 & H_1.

③ تحديد إحصاءة الاختبار ونوع توزيعها.

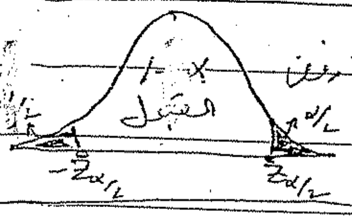
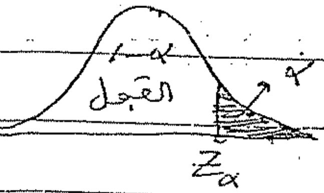
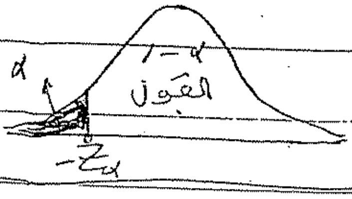
④ حساب قيمة الإحصاءة عندما تكون H_0 صحيحة

⑤ تحديد منطقة الرفض والقبول والقيم الحرجة التي تفصل بينهما.

⑥ اتخاذ القرار المناسب برفض H_0 إذا وقعت القيمة المحسوبة للإحصاءة في منطقة الرفض وقبولها إذا وقعت الإحصاءة المحسوبة في منطقة القبول.

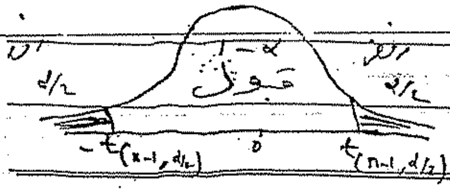
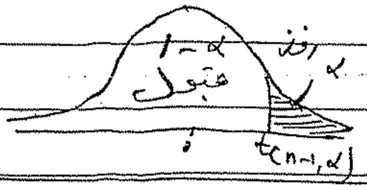
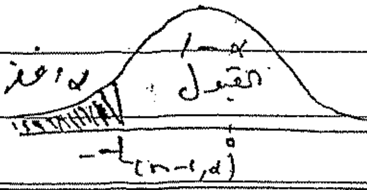
1 اختبارات الفروض حول متوسط المجتمع μ / σ^2 معلومة

وتتلخص في الجدول الآتي:

	الفرضية الإحصائية	إحصاءة الاختبار	القرار
A	$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$Z_c = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$	
B	$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	"	
	$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	"	

※ وإذا كانت σ^2 مجهولة وحجم العينة كبير بدرجة كافية أي $n \geq 30$ فإننا نتبع نفس خطوات الجدول السابق مع استبدال σ بـ s أي أن $Z_c = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}}$

※ أما إذا كانت σ^2 مجهولة + حجم العينة صغير أي $n < 30$ فإن الخطوات ستكون كما في الجدول الآتي:

القرار	إحصاءة الاختبار	الفرضية الإحصائية
	$T=\dfrac{\bar{X}-\mu_0}{s/\sqrt{n}}$	$H_0: \mu=\mu_0$ $H_1: \mu\neq\mu_0$
	"	$H_0: \mu=\mu_0$ $H_1: \mu>\mu_0$
	"	$H_0: \mu=\mu_0$ $H_1: \mu<\mu_0$

مثال / إذا كان متوسط أطوال عينة من 36 طالب بإحدى الكليات هو 68.5 انش بانحراف معياري 7 انش. اختبر الفرض القائل بأن متوسط أطوال جميع الطلبة بالكلية لا يختلف عن 68 انش. استخدم $\alpha=0.05$ كمستوى معنوية؟

الحل

$H_0: \mu=68$ ، σ مجهولة & $n=36>30$

$H_1: \mu\neq68$ ، نستخدم إحصاءة Z مع استبدال σ بـ s وبذلك تكون إحصاءة الاختبار كالتالي:

$$Z_c=\frac{68.5-68}{7/\sqrt{36}}=\boxed{0.43}$$

$\alpha=.05 \Rightarrow \frac{\alpha}{2}=0.025 \Rightarrow Z_{.025}=1.96$

وكون منطقتي الرفض هما كما في الشكل وحيث أن $Z_c=0.43$ فهي واقعة في منطقة القبول نقبل H_0 ونستنتج أن متوسط أطوال الطلبة بالكلية لا يختلف عن 68 انش.

في المثال السابق نفرض أن عدد الطلبة في العينة هو 16

ولإختبار نفس الفرض

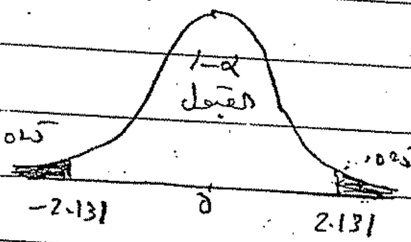
$H_0: \mu = 68$

$H_1: \mu \neq 68$

في هذه الحالة n=16 أقل من 30 و σ مجهولة نستخدم إحصائية T

$$T_c = \frac{68.5 - 68}{7/\sqrt{16}} = \boxed{0.29}$$

ومن جدول t: $t_{(15, .025)} = 2.131$ منطقتي الرفض هما



بما أن Tc = 0.29 تقع في منطقة القبول نقبل H_0 ونستنتج أن متوسط أطوال الطلبة بالكلية لا يختلف عن 68 إنش.

مثال/ مصنع للأدوية يستخدم أسلوب معين لإنتاج نوع معين من الدواء فإذا كان متوسط الإنتاج اليومي بهذا الأسلوب هو 80 وحدة وأرادت إدارة المصنع تغيير أسلوب الإنتاج بغية إحداث زيادة فيه فاستخدمت أسلوبا جديدا لمدة 20 يوما، فكان متوسط الإنتاج خلال هذه الأيام هو 83 وحدة إنتاجية بانحراف معياري 9 وحدات. فهل تدعم نتائج هذه العينة الفرض القائل بأن الأسلوب الجديد أدى إلى زيادة الإنتاج، وذلك باعتبار أن الإنتاج اليومي للمصنع له توزيع طبيعي بمتوسط μ وتباين σ^2 ($\alpha = 0.01$).

الحل

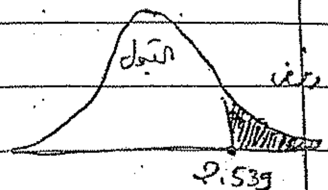
$H_0: \mu = 80$

$H_1: \mu > 80$

n = 20 صغيرة و σ^2 مجهولة ∴ نستخدم إحصائية T

$$T_c = \frac{83-80}{9/\sqrt{20}} = \boxed{1.49}$$

$$t_{(19,\ .01)} = 2.539$$



$\therefore\ T_c < T_{\text{جدولية}}$

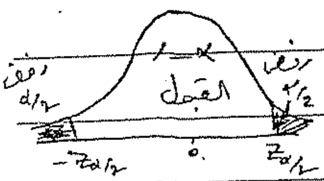
∴ نقبل H_0 أي أن هذه العينة نتائجها لا تدعم أو تؤيد الفرض القائل بأن الأسلوب الجديد أدى إلى الزيادة في الإنتاج.

تلخيص اختبارات الفروض حول المتوسط μ

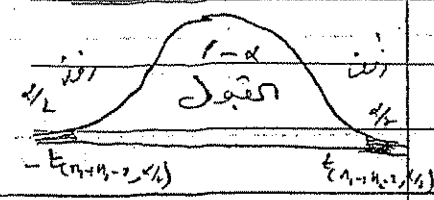
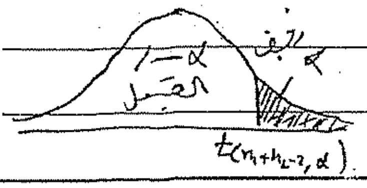
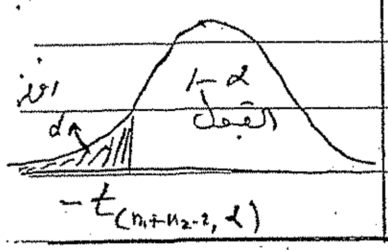
- σ معلومة $\rightarrow$ $Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$
- σ مجهولة & $n \geq 30$ $\rightarrow$ $Z = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$
- $n < 30$ & σ مجهولة $\rightarrow$ $T = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$

II) اختبارات الفروض حول الفرق بين متوسطي مجتمعين μ_1 & μ_2.

A) σ_1^2 & σ_2^2 معلومة، وتتلخص في الجدول الآتي:

الفرضية الإحصائية	إحصاءة الاختبار	القرار
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	$Z_c = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$	
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$	"	
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$	"	

(B) في حالة ($\sigma_1^2=\sigma_2^2=\sigma^2$) مجهولتان و ($30 > n_1+n_2-2$).

الفرضية الإحصائية	إحصاءة الاختبار	القرار
$H_0: \mu_1=\mu_2$ $H_1: \mu_1\neq\mu_2$	$T_c=\dfrac{(\bar x_1-\bar x_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	
$H_0: \mu_1=\mu_2$ $H_1: \mu_1>\mu_2$	"	
$H_0: \mu_1=\mu_2$ $H_1: \mu_1<\mu_2$	"	

حيث $S_p^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$ ويسمى التباين المشترك

وهو مقدر للتباين المجهول σ^2.

في حالة $n_1+n_2-2>30$ تتوزع الإحصاءة T تقريبا التوزيع الطبيعي المعياري أي تصبح

$$Z=\frac{(\bar x_1-\bar x_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \simeq N(0,1)$$

مثال: إذا كان لدينا مجتمعان مستقلان يتبعان التوزيع الطبيعي، وأخذنا عينة عشوائية بحجم 100 مفردة من المجتمع الأول فوجد أن $\bar x_1=68.5$ & $S_1=10$، وعينة أخرى من المجتمع الثاني بحجم 150 مفردة فوجدنا أن $S_2=12$ & $\bar x_2=66.9$، اختبر

فرض العدم $H_0: \mu_1 = \mu_2$ مقابل $H_1: \mu_1 > \mu_2$ عند $\alpha = 1\%$

الحل:

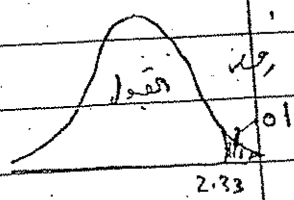
$\bar{x}_1 = 68.5$ $S_1 = 10$ $n_1 = 100$

$\bar{x}_2 = 66.9$ $S_2 = 12$ $n_2 = 150$

$\alpha = .01$ ∵ $n_1, n_2 > 30$ & σ_1^2, σ_2^2 (مجهولتان)

∴ نستخدم إحصاءة الاختبار Z_c حيث

$$Z_c = \frac{68.5 - 66.9}{\sqrt{\frac{(10)^2}{100} + \frac{(12)^2}{150}}} = \frac{1.6}{1.41} = \boxed{1.14}$$



$Z_{.01} = 2.33$

∵ $Z_c = 1.14 < Z_{.01} = 2.33$

∴ واقعة في منطقة القبول ∴ نقبل H_0 ونستنتج أن $\mu_1 = \mu_2$

أي أنه لا يوجد فرق معنوي بين متوسطي الجنسين.

البيانات التالية لعينتين اختيرتا من مجتمعين طبيعيين لهما نفس التباين [$\sigma_1^2 = \sigma_2^2 = \sigma^2$]

رقم العينة	متوسط العينة	حجم العينة	الانحراف المعياري	
1	25	15	6	
2	23	15	8	

اختبر الفرض القائل بأنه لا يوجد اختلاف بين متوسطي العينتين عند $\alpha = .01$

الحل

$H_0: \mu_1 = \mu_2$ ∵ $\sigma_1^2 = \sigma_2^2 = \sigma^2$ مجهولتين

$H_1: \mu_1 \neq \mu_2$ $n_1 + n_2 - 2 < 30$

∴ نستخدم إحصاءة T_c وعليه نحسب أولاً S_p حيث

$$S_p = \sqrt{\frac{14(6)^2 + 14(8)^2}{15+15-2}} = \boxed{7.07}$$

$$\therefore T_c = \frac{(25-23)-0}{7.07\sqrt{\frac{1}{15}+\frac{1}{15}}} = \frac{2}{2.58} = \boxed{0.775}$$

$$\because \alpha = .01 \therefore \frac{\alpha}{2} = .005$$

$$\therefore t_{(28, .005)} = \boxed{2.763}$$

$$-2.763 < 0.775 < 2.763$$

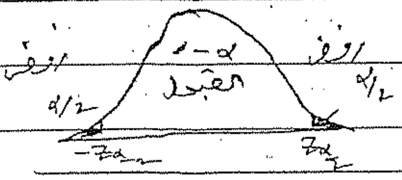
∴ واقعة في منطقة القبول ∴ نقبل H_0 ونستنتج أنه لا يوجد فرق معنوي بين متوسطي العينتين.

III) (ج) اختبارات الفروض حول نسبة المجتمع P

من توزيعات المعاينة أثبتنا أن $\hat{P} \sim N\left(P, \frac{P(1-P)}{n}\right)$

ومنها $Z = \dfrac{\hat{P} - P}{\sqrt{\frac{P(1-P)}{n}}} \sim N(0,1)$

وعليه يمكن تلخيص اختبارات الفروض حول النسبة في الجدول الآتي:

الفرضية الإحصائية	إحصاءة الاختبار	
$H_0: P = P_0$ $H_1: P \neq P_0$	$Z_c = \dfrac{\hat{P} - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}}$	
$H_0: P = P_0$ $H_1: P > P_0$	"	
$H_0: P = P_0$ $H_1: P < P_0$	"	

مثال ١١ يدعي مهندس أن نسبة إصلاح عيب معين في السيارات لا تزيد عن 70% ولاختبار هذا الإدعاء تم إصلاح 76 سيارة من بين مائة سيارة كان بها نفس العيب. فهل هذه العينة تدعم إدعاء المهندس؟ خذ $\alpha = 5\%$ مستوى معنوية

الحل

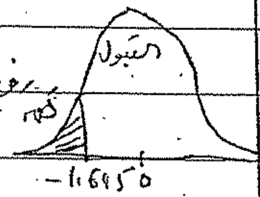
$n = 100 \qquad x = 76 \qquad \therefore \hat{p} = \frac{x}{n} = \frac{76}{100} = .76$

$H_0: P = .70$

$H_1: P < .70 \qquad Z_c = \frac{.76 - .70}{\sqrt{\frac{.70 \times .30}{100}}} = \boxed{1.31}$

$\because Z_c > Z_{table}$

$1.31 > -1.645$



∴ واقعة في منطقة القبول

∴ نقبل H_0 ونستنتج أن نسبة الإصلاح لمثل هذا العيب في السيارات لا تقل عن 70% أي أن هذه العينة لا تدعم إدعاء المهندس.

مثال ٢ إذا كانت نسبة الشفاء من مرض معين بالعلاج التقليدي هي 60% تم اكتشاف علاج جديد تم تجريبه على 150 مريض فتم شفاء منهم 120 فهل هذه العينة تعطي دليلاً كافياً على أن العلاج الجديد أفضل من العلاج التقليدي؟ $\alpha = 2.5\%$

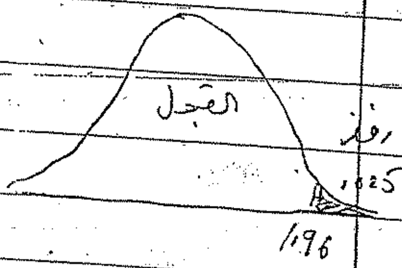
الحل

$\hat{p} = \frac{x}{n} = \frac{120}{150} = .80 \quad , \quad H_0: P = .60$

$H_1: P > .60$

$Z_c = \frac{.80 - .60}{\sqrt{\frac{.60 \times .40}{150}}} = \boxed{5} \quad \& \quad \boxed{Z_{.025} = 1.96}$

$\therefore Z_c = 5 > Z_{.025} = 1.96$



تقع في منطقة الرفض نرفض H_0 ونقبل H_1

ونستنتج أن نسبة الشفاء بالعلاج الجديد

أفضل من العلاج القديم بناءً على

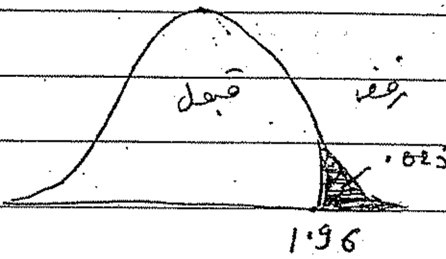
بيانات هذه العينة

∴ $Z_c = 5 > Z_{.025} = 1.96$ (واقع في منطقة الرفض)

∴ نرفض H_0 ونقبل H_1 ونستنتج أن هذه العينة

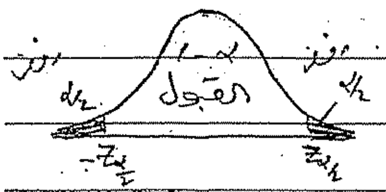
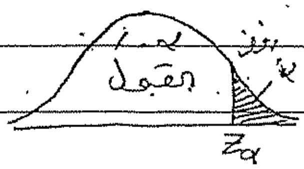
أعطتنا دليلاً كافياً على أن نسبة الشفاء بالعلاج الجديد

أفضل من نسبة الشفاء بالعلاج التقليدي.



IV. اختبارات الفروض حول الفرق بين نسبتي مجتمعين P_1 & P_2

ويمكن تلخيص ذلك في الجدول الآتي :-

الفرضية الإحصائية	إحصاءة الاختبار	القرار
$H_0: P_1 = P_2$ $H_1: P_1 \neq P_2$	$Z_c = \dfrac{\hat{P}_1 - \hat{P}_2}{\sqrt{\bar{P}_0(1-\bar{P}_0)\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$	
$H_0: P_1 = P_2$ $H_1: P_1 > P_2$	"	
$H_0: P_1 = P_2$ $H_1: P_1 < P_2$	"	

حيث $\bar{P}_0 = \dfrac{x_1 + x_2}{n_1 + n_2}$ or $\bar{P}_0 = \dfrac{n_1 \hat{P}_1 + n_2 \hat{P}_2}{n_1 + n_2}$

مثال: مصنع أدوية ينتج نوعين من الدواء لتخفيف آلام الصداع، فإذا جرب النوع الأول على 100 فرد من يعانون من آلام الصداع فنجح في تخفيف الآلام 90 منهم، وجرب النوع الثاني على 150 مريض فلم يخفف الآلام 20 منهم، اختبر الفرض القائل بأن النوعين من الدواء لهما نفس التأثير؟ عند $\alpha = .05$

الحل

$n_1 = 100 \qquad x_1 = 90 \Rightarrow \hat{p}_1 = \frac{90}{100} = .9$

$n_2 = 150 \qquad x_2 = 130 \Rightarrow \hat{p}_2 = \frac{130}{150} = 0.87$

$$P_0 = \frac{x_1 + x_2}{n_1 + n_2} = \frac{90 + 130}{100 + 150} = \frac{220}{250} = 0.88$$

$$\therefore Z_c = \frac{.90 - 0.87}{\sqrt{.88 \times .12\left[\frac{1}{100} + \frac{1}{150}\right]}} = \frac{.03}{.042} = \boxed{0.71}$$

$H_0: P_1 = P_2 = P_0$

$H_1: P_1 \neq P_2$

$Z_{\frac{\alpha}{2}} = Z_{.025} = \boxed{1.96}$

$\therefore Z_c = 0.71$ واقعة في منطقة القبول

$\therefore$ نقبل H_0 ونستنتج أن الدوائين لهما نفس التأثير.

تمرين: مصنع الطوب الاسمنتي تم تطوير طريقة جديدة لصناعة الطوب الاسمنتي ولإثبات أفضليتها على الطريقة القديمة أخذت عينة حجمها 200 من الطريقة القديمة فوجد بها 16 طوبة مكسورة بينما وجدت 24 طوبة مكسورة في عينة من 300 طوبة أنتجت بالطريقة الجديدة، اختبر أفضلية الطريقة الجديدة عند مستوى المعنوية 5% ؟

الارتباط والانحدار

الارتباط: هو معيار لقياس قوة واتجاه العلاقة بين المتغيرات، وقد يكون قويا أو ضعيفا أو معدوما تبعا للعلاقة. وقد يكون موجبا (طرديا) أو سالبا (عكسيا) حسب اتجاه العلاقة.

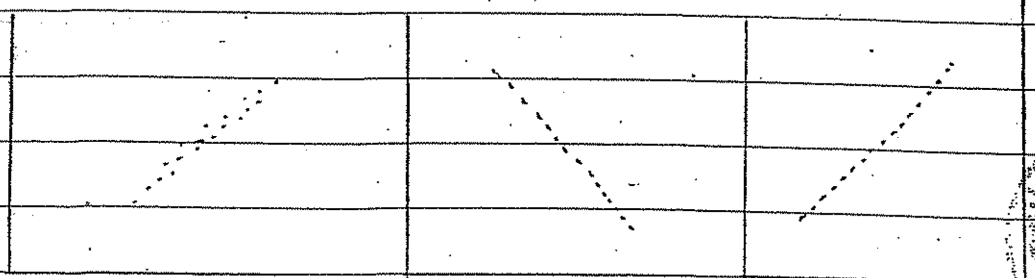
أنواع الارتباط:

1- الارتباط البسيط: وهو الذي يبحث في العلاقة بين متغيرين اثنين فقط، احدهما تابع والاخر مستقل.

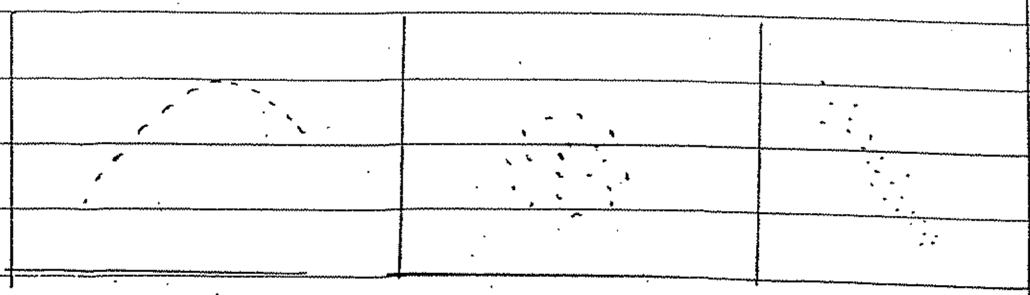
2- الارتباط المتعدد: ويبحث في العلاقة بين أكثر من متغيرين.

3- الارتباط الجزئي: ويبحث في العلاقة بين مجموعة من المتغيرات مع عزل تأثير بعض المتغيرات الأخرى.

الارتباط الخطي البسيط: يمكن أخذ فكرة عامة حول العلاقة ونوعيتها من خلال تفحص الشكل الانتشاري لقيم المتغيرات تمثل الشكل



ارتباط طردي تام (A) — عكسي تام (B) — (C) طردي قوي



(D) عكسي قويا — (E) منعدم — (F) غير خطي

ويعتبر الارتباط الخطي الأكثر استعمالاً، وذلك نظراً لأن معظم العلاقات الغير خطية يمكن تقريبها بشكل فرضي إلى علاقة خطية. وعموماً فالارتباط الخطي البسيط يقيس قوة واتجاه العلاقة الخطية بين متغيرين، وعادة ما يقاس بمعامل يسمى معامل ارتباط بيرسون أو معامل ارتباط العينة ويرمز له بالرمز "r" حيث

$$r = \frac{S_{xy}}{S_x S_y}$$

حيث

$$S_{xy} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{n-1}$$

ويسمى التغاير بين x و y.

S_x : الانحراف المعياري لقيم x

S_y : y = = =

أي أن

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2}\sqrt{\sum(y_i-\bar{y})^2}}$$

ويمكن كتابتها بصيغة أسهل في الحساب

$$r = \frac{\cancel{n}\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sqrt{\left[\sum x_i^2 - n\bar{x}^2\right]\left[\sum y_i^2 - n\bar{y}^2\right]}}$$

$$r = \frac{n\sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{\left[n\sum x_i^2 - (\sum x_i)^2\right]\left[n\sum y_i^2 - (\sum y_i)^2\right]}}$$

ملاحظة: قيمة معامل الارتباط وإشارته لا تتغير إذا قمنا بتبديل أحد المتغيرين بالآخر.

خواص معامل الارتباط:

1- إذا كانت العلاقة الخطية بين المتغيرين منعدمة فإن قيمة معامل الارتباط تساوي صفراً أي أن $\boxed{r=0}$

2- إذا كانت إشارة معامل الارتباط «سالبة» تكون العلاقة عكسية أي كلما زادت قيمة أحد المتغيرين قلّت قيمة المتغير الآخر والعكس صحيح.

3- إذا كانت إشارة معامل الارتباط «موجبة» تكون العلاقة طردية أي كلما زادت قيمة أحد المتغيرين زادت قيمة المتغير الآخر.

4- إذا كانت $\boxed{r=+1}$ (طردية تامة) أما إذا كانت $\boxed{r=-1}$ (عكسية تامة)

5- قيمة معامل الارتباط تنحصر بين (-1) و (+1) أي أن

$$\boxed{-1 \le r \le +1}$$

مثال: أوجد معامل الارتباط بين المتغيرين X و Y في الجدول الآتي:

X	1	3	4	6	6	20
Y	10	2	7	6	5	30
XY	10	6	28	36	30	110
X^2	1	9	16	36	36	98
Y^2	100	4	49	36	25	214

$$\therefore r = \frac{5(110)-(20)(30)}{\sqrt{[5(98)-(20)^2][5(214)-(30)^2]}} = \frac{-50}{\sqrt{(90)(170)}} = \boxed{-.404}$$

عكسية ضعيفة

مثال: لدراسة العلاقة بين نسبة المادة الداخلة في تركيب معدن معين (X) وكمية صلابته (Y)، أخذت عينة عشوائية مكونة من 10 قضبان من ذلك المعدن وحسبت نسبة المادة في المعدن ودرجة الصلابة في كل قضيب فكانت كالتالي:

											$\sum$
نسبة المادة X	5	2	8	12	4	6	8	2	9	4	60
درجة الصلابة Y	16	10	22	25	15	18	24	9	20	15	174
XY	80	20	176	300	60	108	192	18	180	60	1194
X^2	25	04	64	144	16	36	64	04	81	16	454
Y^2	256	100	484	625	225	324	576	81	400	225	3296

$$\therefore r = \frac{(10)(1194) - (60)(174)}{\sqrt{[(10)(454) - (60)^2][(10)(3296) - (174)^2]}} = \boxed{0.94}$$

∴ يمكن القول من خلال هذه العينة أن العلاقة بين نسبة المادة الداخلة في تركيب المعدن ودرجة صلابته هي علاقة طردية قوية جداً.

الانحدار الخطي البسيط

من خلال وجود علاقة خطية بين المتغيرين، فإنه علينا إيجاد المعادلة الرياضية التي تربط بينهما مع تحديد أي المتغيرين يعتمد على الآخر ويتأثر به، ويسمى المتغير التابع، أما الآخر فيسمى المتغير المستقل، وفي الغالب توجد علاقة سببية بين المتغيرين يعرف أيهما المستقل وأيهما التابع.

فعند دراسة العلاقة بين سرعة السيارة واستهلاك الوقود نجد أن استهلاك الوقود يعتمد على سرعة السيارة، وليس العكس، فإن سرعة السيارة هي المتغير المستقل وكمية الوقود

المستقل هو المتغير التابع له.

معادلة انحدار Y على X: تسمى المعادلة الرياضية التي تربط بين المتغير التابع Y بالمتغير المستقل X بمعادلة انحدار Y على X. ولأن العلاقة بينهما خطية فتكون على صورة معادلة خط مستقيم من الدرجة الأولى:-

$$\boxed{y_i = a + bx_i}$$

حيث a و b ثابتان يجب تقديرهما من قيم العينة العشوائية بطريقة تسمى طريقة المربعات الصغرى التي تعتمد على اختيار الخط الأمثل الذي يمثل نقاط العينة في شكلها الإنتشاري، والذي يجعل قيمة مجموع مربعات الفروق (الأخطاء) بين القيم الحقيقية y_i والقيم المقدرة $\hat{y}_i$ عن طريق خط الإنحدار أقل ما يمكن. ويفرض أن قيمتي a و b المقدرتان بطريقة المربعات الصغرى هما: $\hat{a}$ و $\hat{b}$ وبذلك تكون معادلة انحدار Y على X هي:-

$$\hat{Y} = \hat{a} + \hat{b}X$$

والخطأ بين القيمة الحقيقية y_i والمقدرة $\hat{y}_i$ هو

$$e_i = y_i - \hat{y}_i \qquad i = 1, 2, \ldots, n$$

$$= y_i - (\hat{a} + \hat{b}x_i)$$

$$e_i = y_i - \hat{a} - \hat{b}x_i$$

ومجموع مربعات الأخطاء هو

$$S = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{a} - \hat{b}x_i)^2$$

وبالإستعانة بحساب التفاضل فإنه علينا إيجاد قيمتي $\hat{a}$ و b اللتان تجعلان المجموع S في حالة نهاية صغرى، وذلك بإيجاد المشتقات الجزئية للدالة S بالنسبة لكل من $\hat{a}$ و $\hat{b}$ وحل المعادلتين الناتجتين بعد مساواتهما بالصفر فتعطي

مثال 11

$A, B, C, D, E, F \rightarrow 6$

$= 6 \times 5 \times 4$

$1 - P(X < 1) = 0.9999$

$= 0.9999$