وزارة التعليم العالي والبحث العلمي

جامعة الجيلالي بونعامة خميس مليانة

كلية العلوم اجتماعية

قسم علم الاجتماع

مقياس احصاء استدلالي ورياضي

السنة الثانية علم الاجتماع

الاستاذة : زرقي عائشة

السنة الجامعية 2022/2021

المحاضرة الاولى: المعالجة الاحصائية للمقاييس الاسمية

اختبار كاي تربيع x^2

تمهيد

يعتبر إختبار مربع كاي من الاختبارات اللابارامترية الأكثر استخداما في البحوث النفسية والتربوية والاجتماعية، وهو يستعمل في حالة البيانات الإسمية(مستوى قياس إسمي/تصنيفي)،حيث يصنف الأفراد أو الأشياء أو الإستجابات في فئات أو أقسام مختلفة. ومثال ذلك الإستجابة على بنود استبيان يحتوي فقرات تتطلب الاجابة عن كل فقرة بديلا من 3(موافق، محايد، معارض

- إن هدف اختبار مربع كاي هو المقارنة بين التوزيع التكراري المشاهد والتوزيع التكراري المتوقع لإحدى العينات إذا افترضنا وجود توزيع احتمالي للمجتمع أو التحقق ما إذا كانت هناك فروق دالة إحصائيا بين التكرارات المشاهدة والتكرارات المتوقعة لاستجابات مجموعة من الأفراد.

يطبق اختبار مربع كاي في حالة وجود متغير نوعي واحد ويسمى اختبار حسن المطابقة، أو في حالة وجود متغيرين أو أكثر ويسمى كا2 للاستقلالية.

<u>شروط تطبيق كا2</u>

- البيانات نوعية
- مستوى القياس إسمي
- أن تكون العينة منتقاة عشوائيا من المجتمع المستهدف.
- أن يكون حجم العينة أكبر من 30وهذا لأن اختبار كا2 يتأثر تأثرا واضحا بالعينات الصغيرة .
- أن لا يقل أي تكرار متوقع في كل خلية من خلايا الجدول عن5 وذلك لأنه إذا كانت التكرارات المتوقعة صغيرة جدا فإن توزيع معاينات كا2 لن تكون ممثلة تمثيلا كافيا لتوزيع المتغير موضع البحث.

ويعطى كا^2 بالمعادلة التالية:

- **معادلة اختبار X^2:**

$$\mathbf{x^2 = \Sigma \frac{(f_o - f_t)^2}{f_t}}$$

حيث:

f_oيمثل التكرار المشاهد(fréquences observées)

f_t يمثل التكرار النظري (fréquences théoriques)

$$(كا^2) = \frac{(\text{التكرار المشاهد} - \text{التكرار النظري})^2}{\text{التكرار النظري}}$$

أو بالرموز كالآتي:

$$(كا^2) = \frac{(ك_{ش} - ك_{ن})^2}{ك_{ن}}$$

إذ أن :

$كا^2$ = مجموع قيم مربع كاي المحسوبة لكل الخلايا.

ك ش= التكرار المشاهد .

ك ن = التكرارات النظرية (المتوقعة).

هذا ويتم تقويم قيمة (كا2) المحسوبة بالرجوع إلى الجداول الإحصائية الخاصة بالقيم الجدولية لمربع كاي عند درجات حرية تتوقف على عدد الخلايا أو فئات التصنيف في التجربة.

خطوات حساب قيمة مربع كاي (كا2) لعامل واحد :

يمكن حساب قيمة مربع كاي المحسوبة من خلال:

*تحديد الفئات المختلفة للتصنيف بحيث لا تقل عن (2 فئات).

*توزيع المشاهدات على هذه الفئات , بحيث توضع كل مشاهدة في فئة واحدة فقط من فئات التوزيع.

*تحديد التكرارات المتوقعة (النظرية) في كل فئة من فئات التصنيف, باستخدام القانون التالي:

مج العينة الكلي

التكرار النظري = ـــــــ

فئات التصنيف

*نحسب قيمة مربع كاي لكل خلية بتطبيق المعادلة التالية :

(ك ش – كن)2

(كا2)=ـــــ

ك ن

$$\mathbf{x^2 = \Sigma \frac{(f_o - f_t)^2}{f_t}}$$

- ✓ *نجمع قيم مربع كاي الخاصة بكل الفئات لنحصل على النتيجة النهائية لاختبار (كا2).
- ✓ *نحدد درجة الحرية :

درجة الحرية =عدد الخلايا (فئات التصنيف) – 1.

- ✓ *نحدد قيمة (كا2) الجدولية عند درجة الحرية المحددة ونسبة الخطأ المختارة .
- ✓ *نقارن بين قيمة (كا2) المحسوبة بالقيمة الجدولية فإذا كانت القيمة المحسوبة تساوي أو اكبر من القيمة الجدولية تكون التكرارات المشاهدة مختلفة عن التكرارات المتوقعة وتكون قيم (كا2) المحسوبة دالة إحصائيا أي توجد فروق معنوية بين التكرارات المشاهدة والتكرارات النظرية , أما إذا كانت قيمة (كا2) المحسوبة اصغر من قيمتها الجدولية فتعني هذه الحالة أن التكرارات المشاهدة لا تختلف عن التكرارات أي لا توجد فروق معنوية بين التكرارات المشاهدة والتكرارات النظرية

مثال /

وزع استبيان على عينة من الطلاب بلغ عددهم (60 طالبا) لبيان ممارسة الطلبة للنشاط الرياضي فظهر أن (32) طالب يمارسون الأنشطة الرياضية و(28) طالب لا يمارسون أي نشاط رياضي , المطلوب بيان الفرق بين الممارسين وغير الممارسين للأنشطة الرياضية.

الحل /

$$(كا^2) = \frac{(\text{التكرار المشاهد} - \text{التكرارات المتوقع})^2}{\text{التكرارات المتوقع}}$$

$$(كا^2) = \frac{(ك_{ش} - ك_{م})^2}{ك_{م}}$$

f_o التكرار المشاهد للممارسين = 32 طالب

التكرار المشاهد لغير الممارسين = 28 طالب

$$ft \text{ التكرار النظري} = \frac{\text{مج العينة الكلي}}{\text{فئات التصنيف}} = \frac{60}{2} = 30$$

نكمل الحل باستخدام الجدول التالي

ئات التصنيف	لتكرار لمشاهد ($ك_ش$) f_o	لتكرار النظري ($ك_م$) ft	($ك_ش$-$ك_م$) $(f_o - f_t$	($ك_ش$- $ك_م$)^2 $\frac{(f_o - f_t)^2}{}$	($ك_ش$- $ك_م$)^2 $\frac{}{ك_م}$ $\Sigma\frac{(f_o - f_t)^2}{f_t}$	يمة ($كا^2$) لمحسوبة
ممارس	32	30	2	4	0.13=30÷4	=0.13+0.13
لا يمارس	28	30	2-	4	0.13=30÷4	0.26

نجد درجة الحرية = عدد الخلايا (فئات التصنيف) -1

1 = 1-2 =

إذن قيمة ($كا^2$) المحسوبة تساوي (0.26) وهي قيمة اصغر من قيمتها الجدولية البالغة (3.84) عند درجة حرية (1) ومستوى دلالة (0.05) وهذا يعني عدم وجود فروق معنوية بين الطلبة الممارسين وغير الممارسين للأنشطة الرياضية

تمرين:

قام احد الباحثين باستطلاع رأي لعينة من الطلاب في إحدى الجامعات لمعرفة آرائهم حول اعتماد درس التربية الرياضية في الجامعات وكانت العينة تتضمن (500)طالب من طلبة المرحلة الأولى و(350)طالب من المرحلة الثانية والثالثة و(250) طالب من المرحلة الرابعة وكانت النتائج وبدائل الإجابة كما مبين في الجدول التالي :

مرحلة	دائل الإجابة			مجموع للصف
	مادة اختيارية	مادة إجبارية	م أكون رأي	
لأولى	167	272	61	500
ثانية والثالثة	149	157	44	350
رابعة	115	105	30	250

مجموع العمود	431	534	135	1100

المطلوب اختبار الفرض الصفري الذي يقرر أنه لا تأثير للمستوى الدراسي على آراء الطلبة حول درس التربية الرياضية في الجامعات

المحاضرة الثانية : التنبؤ

كيف نستطيع التنبؤ بظاهرة معينة من خلال معرفة ما له علاقة بهذه الظاهرة.

(إذاكان كذا ، سيكون كذا)

التنبؤ علاقة بين متغيرين $[\ y = f(x)\ ]$ $x\ ,\ y$ من خلال f وهي دالة تربط بين متغيرين يفترض الباحث أن أحدها تابع والآخر مستقل

والتنبؤ هو معرفة أحد المتغيرين بمعلومية الآخر ، والمتغير المطلوب معرفته يسمى (المتغير التابع) ، والمتغير المعلوم يسمى (المتغير المستقل)

وهذا هو أبسط وأسهل أنواع التنبؤ $[\ y = f(x)\ ]$، فالبساطة جاءت من عدد المتغيرات فهو أقل ما يكون. وقد تكون أكثر من متغير $[\ y = f(x_1, x_2, x_3 \dots \dots \dots \dots . x_n\ ]$ وهذا أكثر واقعية لأن الظاهرة تتأثر بعد متغيرات وليس متغيرا واحدا. فالتحصيل لا يتأثر بالذكاء فقط بل أيضا بالبيئة وبالمثابرة وبتعلم الوالدين .....الخ

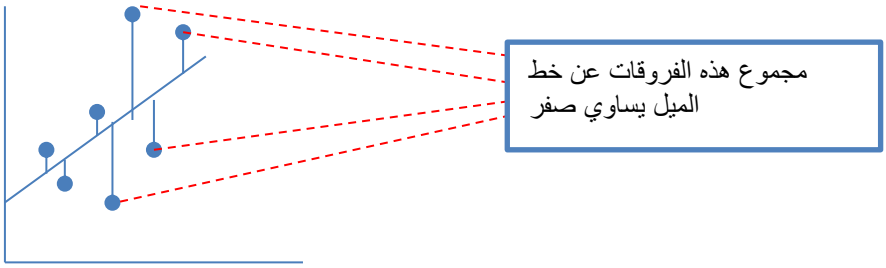
فالسهولة هنا جاءت من المعادلة الرياضية فهذه الدالة فقط عن العلاقة الخطية ، ولكن هناك صور للعلاقات غير خطية ، فهناك العلاقات المنحنية وعلاقات تأخذ شكل حرف S وعلاقات عكسية وعلاقات لوغارتمية وعلاقات أسية.

فيجب معرفة طبيعة العلاقة بين الظاهر هل هي خطية أم لا ، ويكون ذلك من خلال الرسم للبيانات. وهناك طرق أخرى بالحسابات لكنها أصعب.

أسلوب تحليل الانحدار

الأساس العلمي الذي بني عليه أسلوب تحليل الانحدار يعتمد على طريقة المربعات الصغرى (OLS) Ordinary Least Squares وتتلخص هذه الطريقة في أن تجعل مجموع مربعات الأخطاء أقل ما يمكن.

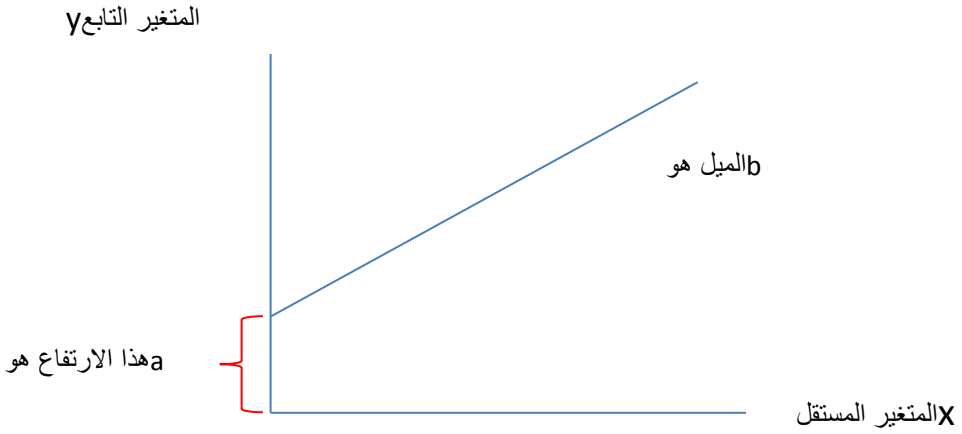
الفرق بين التقدير والمعلمة يسمى الخطأ والمجموع يساوي صفر $(\mu - \bar{x}) = 0$



تفترض أن معادلة الانحدار الخطي البسيط بين متغيرين أحدهما تابع والآخر مستقل على الصورة التالية:

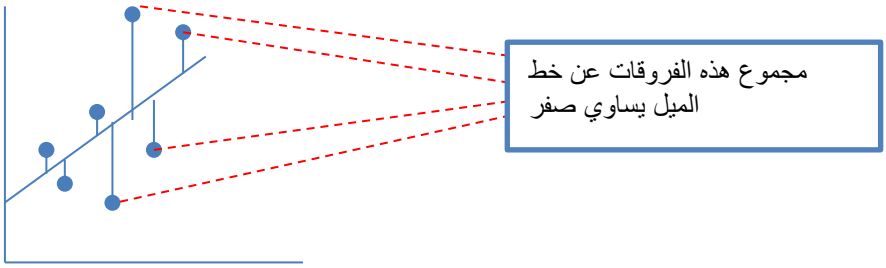
$$Y = a + bx + \epsilon$$

المتغير التابع | الجزء المقطوع من المحور الرأسي (ثابت) | ميل الخط المستقيم على المحور الأفقي | المتغير المستقل | الخطأ العشوائي



$$y = a + bx + \epsilon$$

يبقى أن نعرف كيف نقدر (a) و (b) لذلك يجب أن يختفي الخطأ العشوائي(ϵ) من المعادلة ، وذلك من خلال طريقة المربعات الصغرى التي تجعل مجموع مربعات الأخطاء أقل ما يمكن (أي صفر) كما في الشكل التالي:



وتصبح المعادلة هكذا: $\hat{Y} = \hat{a} + \hat{b}x$

و نلاحظ أنxليس عليها علامة ^ لأنها معلومة في الأصل

حيث أن ^ عبارة عن : (معامل الارتباط × $\frac{\text{الانحراف المعياري y}}{\text{الانحراف المعياري x}}$) وتحسب بالمعادلة التالية:

$$\hat{b} = r \times \frac{S_y}{S_x}$$

وتدل (b) على كمية الزيادة في (y) إذا زادت (x) بوحدة واحدة.

كما أن $\hat{a}$ عبارة عن الثابت في المعادلة ، تحسب عن طريق المعادلة التالية:

$$\hat{a} = \bar{y} - \hat{b}\bar{x}$$

حيث$(\bar{x})$ متوسط المتغير المستقل و $(\bar{y})$ متوسط المتغير التابع

=====

معامل التحديدr^2(التفسير) : يتراوح بين (صفر و 1) ، دائما موجب فهو مربع معامل الارتباط.

انه نسبة مساهمة المتغير المستقيل في تفسير التغيرات التي تحدث في المتغير التابع

تمرين1 :

- أوجد معامل الارتباط بين المتغيرين (x) و(Y).
- احسب معامل التحديد واشرح معناه
- قدّر معادلة انحدار (y) على (x) [معادلة التنبؤ] $\hat{} = \hat{a} + \hat{b}(x)$
- ارسم الشكل الانتشاري للعلاقة بين المتغيرين، مع رسم خط ميل الانحدار.

اسم الطالب	x	y	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
محمد	1	2					
فهد	2	4					
سعد	3	6					
خالد	4	8					
سعود	5	10					
المجموعΣ	15	30					

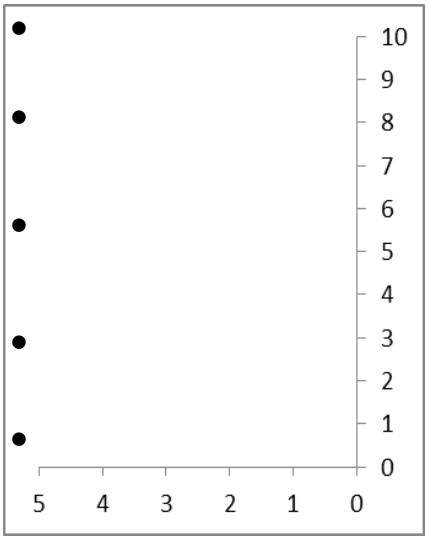
$$\bar{x} = \frac{\sum x}{n} = - =$$

$$\bar{y} = \frac{\sum y}{n} = - =$$

$$S_x = \sqrt{\frac{\sum (x-\bar{x})^2}{n}} = \sqrt{-} = \sqrt{\quad} =$$

$$S_y = \sqrt{\frac{\sum (y-\bar{y})^2}{n}} = \sqrt{-} = \sqrt{\quad} =$$

$$r = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\sum (x-\bar{x})^2 \sum (y-\bar{y})^2}} = \frac{\quad}{\sqrt{\quad} \times} = \frac{\quad}{\sqrt{\quad}} = - =$$



- **التعليق** : توجد علاقة ارتباط ..... بين عدد ساعات المذاكرة والدرجة التي يحصل عليها الطالب

- **معامل التحديد r^2 يساوي** .....أي أن المتغير (x) يفسر التغيرات التي يحدث في المتغير(y) بنسبة .....%

- $\hat{b} = r \times \frac{S_y}{S_x} = \quad \times - = \quad \times \quad =$

- $\hat{a} = \bar{y} - (\hat{b} \times \bar{x}) = \quad -(\quad \times \quad) =$

إذن معادلة انحدار(y)على(x) في هذا المثال هي: $\hat{y} = \hat{a} + \hat{\ } (x)$

$$\hat{y} = \quad + \quad (\quad)$$

تطبيق

- **تنبأبالقيم التي ستكون للمتغير (y) في حالات قيم (X) التالية ، ثم حدد مكانها على الشكل الانتشاري:**

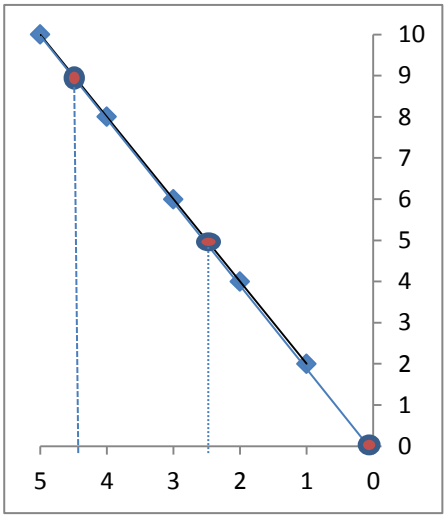
x=4.5	x=2.5	x= 0
$\hat{y} = \quad + \quad (\quad) =$	$\hat{y} = \quad + \quad (\quad) =$	$\hat{y} = \quad + \quad (\quad) =$

حل تمرين1 :

- أوجد معامل الارتباط بين المتغيرين (x) و(Y).
- احسب معامل التحديد واشرح معناه
- قدّر معادلة انحدار (y) على (x) [معادلة التنبؤ] $\hat{\ } = \hat{a} + \hat{b}(x)$
- ارسم الشكل الانتشاري للعلاقة بين المتغيرين، مع رسم خط ميل الانحدار.

اسم الطالب	x	y	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
محمد	1	2	-2	-4	4	16	8
فهد	2	4	-1	-2	1	4	2
سعد	3	6	0	0	0	0	0
خالد	4	8	1	2	1	4	2

سعود	5	10	2	4	4	16	8
المجموع∑	15	30			10	40	20



$$\bar{x} = \frac{\sum x}{n} = \frac{15}{5} = 3$$

$$\bar{y} = \frac{\sum y}{n} = \frac{30}{5} = 6$$

$$S_x = \sqrt{\frac{\sum(x-\bar{x})^2}{n}} = \sqrt{\frac{10}{5}} = \sqrt{2} = 1.41$$

$$S_y = \sqrt{\frac{\sum(y-\bar{y})^2}{n}} = \sqrt{\frac{40}{5}} = \sqrt{8} = 2.82$$

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}} = \frac{20}{\sqrt{10\times40}} = \frac{20}{\sqrt{40}} = \frac{20}{20} = 1.00$$

- **التعليق** : توجد علاقة ارتباط تامة بين عدد ساعات المذاكرة والدرجة التي يحصل عليها الطالب
- **معامل التحديد r^2 يساوي 1.00** أي أن المتغير (x) يفسر التغيرات التي يحدث في المتغير(y) بنسبة 100%
- $\hat{b} = r \times \frac{S_y}{S_x} = 1.00 \times \frac{2.82}{1.41} = 1.00 \times 2.00 = 2.00$
- $\hat{a} = \bar{y} - \hat{b}\bar{x} = 6 - 2(3) = .0$
- إذن معادلة انحدار(y)على(x) في هذا المثال هي: $\hat{y} = 0 + 2(x)$

تطبيق

- **قدّر القيم التي ستكون للمتغير (y) في حالات قيم (X) التالية ، ثم حدد مكانها على الشكل الانتشاري:**

x=4.5	x=2.5	x= 0
$\hat{y} = 0 + 2(4.5) = 9$	$\hat{y} = 0 + 2(2.5) = 5$	$\hat{y} = 0 + 2\,(0) = 0$

التوزيعات الاحصائية هي توزيعات احتمالية ، والاحتمالات هي موضوعات رياضية تدرس سلوك ظاهرة معينة بالأرقام وتحدد أمكانية حدوثها $0 \leq P(A) \leq 1$

وقد قدم علم الاحصاء العديد من التوزيعات الاحصائية للمتغيرات العشوائية التي تسهل التعامل معها ومتابعة سلوك الظاهرة أو المتغير وحساب الاحتمالات المتغيرة.

المتغيرات العشوائية تنقسم إلى قسمين/

أولا/ متغيرات عشوائية متقطعة:

وهي المتغيرات التي تأخذ قيما منفصلة بينها فراغات غير متصلة ، وهذا النوع له توزيعات إحصائية هامة منها:

1) توزيع برونللي.
2) توزيع ذي الحدين.
3) توزيع بواسون.
4) التوزيع الهندسي.
5) التوزيع الهندسي الزائدي.
6) توزيع ذي الحدين السالب.
7) التوزيع المنتظم.

ثانيا/ متغيرات عشوائية متصلة:

وهي المتغيرات التي تأخذ قيما متواصلة لا تنقطع ولا يوجد بينها فجوات أو فواصل مثل : الوزن ، والعمر، والطول، والدخل، والزمن، والمسافة، الدرجات .... وهذه لها توزيعات إحصائية متصلة مثل:

1) التوزيع الطبيعي.
2) توزيع $\boldsymbol{t}$.
3) توزيع $\boldsymbol{f}$.
4) توزيع $\boldsymbol{x^2}$.
5) التوزيع المنتظم المتصل.

6) توزيع بيتا.
7) توزيع جاما.
8) توزيع كوشي.

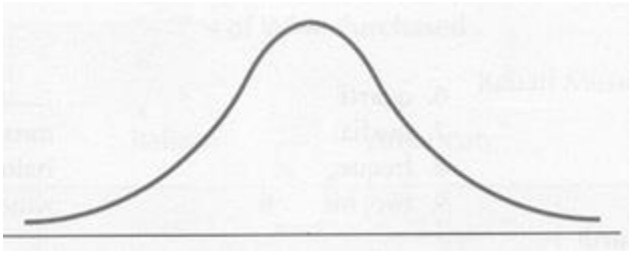
التوزيع الطبيعي

التوزيع الطبيعي أهم التوزيعات الاحصائية على الإطلاق لسببين:

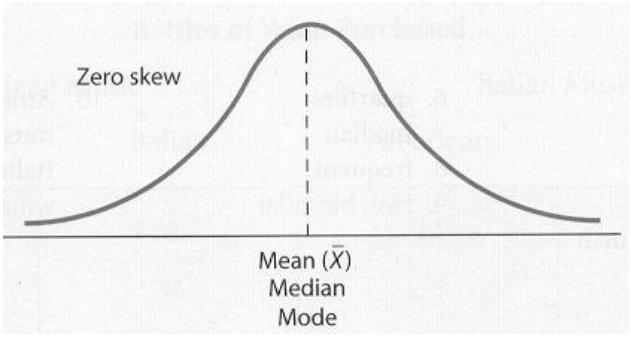
- السبب الأول : أن أغلب الظواهر الحياتية تتبع هذا التوزيع (الطول، الوزن ، العمر، الذكاء ....)
- السبب الثاني: بيانات الظواهر التي لا تتبع التوزيع الطبيعي أي التي تتبع توزيعات أخرى عند زيادة حجمها فإنها تتوزع طبيعيا.

خصائص التوزيع الطبيعي:

1- منحنى التوزيع الطبيعي يأخذ شكل الجرس.

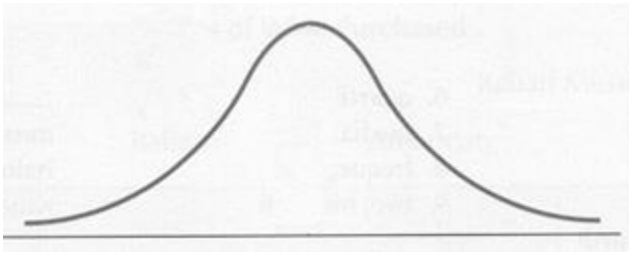


2- منحنى التوزيع الطبيعي متماثل حول المتوسط الحسابي ، وهذا التماثل يعني أن: معامل الالتواء يساوي صفر.
3- قمة منحنى التوزيع الطبيعي قمة معتدلة ليست مرتفعة مدببة، وليست منخفضة مفرطحة. وهذا معناه أن معامل التفرطح يساوي 3
4- في التوزيع الطبيعي يتساوى المتوسط والوسيط والمنوال.



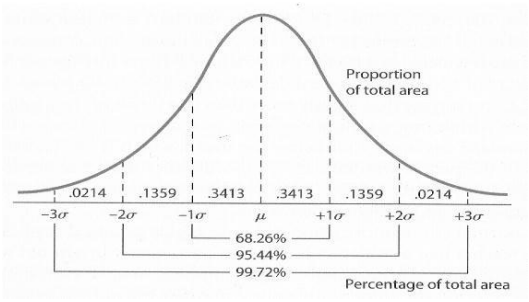
5- التوزيع الطبيعي يعتمد على معلمتين همنا المتوسط μ والتباين σ^2.

6- طرفي منحنى التوزيع الاعتدالي ممتدان إلى ما لا نهاية ولا يقاطعان ولا يمسان المحور الأفقي.



7- المساحة تحت المنحنى الطبيعي محسوبة كما يلي:

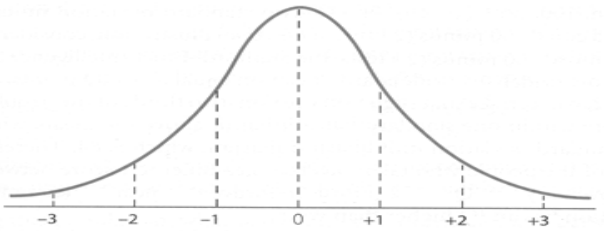
- المساحة المحصورة تحت المنحنى الطبيعي بين $(\mu - 1\sigma$ و$\mu + 1\sigma)$ تساوي 68%
- المساحة المحصورة تحت المنحنى الطبيعي بين $(\mu - 2\sigma$ و$\mu + 2\sigma)$ تساوي 95%
- المساحة المحصورة تحت المنحنى الطبيعي بين $(\mu - 3\sigma$ و$\mu + 3\sigma)$ تساوي 99%
- المساحة الكلية المحصورة تحت المنحنى الطبيعي تساوي واحد صحيح



8- يمكن تحويل منحنى التوزيع الطبيعي إلى منحنى توزيع معياري وذلك بتحويل قيم المتغير العشوائي الأصلية إلى درجات معيارية من خلال المعادلة التالية:

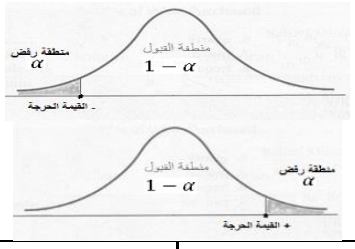
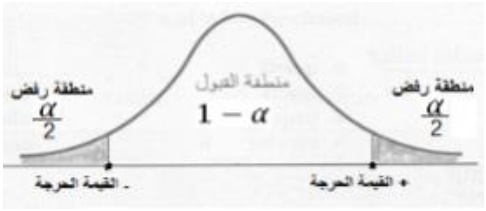
الدرجة المعيارية $\mathbf{Z} = \frac{x-\mu}{\sigma}$

9- التوزيع الطبيعي المعياري له نفس الخصائص ما عدا كل ما يتعلق بالمتوسط μ والتباين σ^2 حيث أنهما أصبحا (صفر) و (واحد).



10- متوسط التوزيع الطبيعي المعياري (صفر) وتباينه (1) ، ولذلك يمكن تحويل أي توزيع طبيعي إلى توزيع طبيعي معياري له جداول إحصائية واحدة .

11- القيم الجدولية من جدول التوزيع الطبيعي المعياري الهامة في اختبار(z) يمكن تلخيصها في أربع قيم كما يلي:

<table>
<tr><td colspan="2">ذيل واحد
 
</td><td colspan="2">ذيلين
 
</td></tr>
<tr><td>$\alpha = 0.05$</td><td>$\alpha = 0.01$</td><td>$\frac{\alpha}{2}$ $\alpha = 0.05$
$= 0.025$</td><td>$\frac{\alpha}{2}$ $\alpha = 0.01$
$= 0.005$</td></tr>
<tr><td>1.64</td><td>2.33</td><td>1.96</td><td>2.58</td></tr>
</table>

توزيع ذي الحدين: يعتبر هذا التوزيع من التوزيعات الاحتمالية المنفصلة، ويستخدم في الظواهر التي تكون النتائج الممكنة لها واحدة من اثنتين احداهما تسمى نجاحا والثانية تسمى فشلا،

والمقصود بالنجاح والفشل هنا هو حدوث الحادث(نجاح) وعدم حدوثه أو تحققه(فشل).وعلى سبيل المثال: نجاح/رسوب، غياب/حضور، وصول/عدم وصول.

وتحدث باحتمال ل بحيث : ل= 1−ح . وعند تكرار التجربة عددا من المرات فإننا نحصل في كل مرة إما على حالة نجاح باحتمال ح أو حالة فشل باحتمال ل.

والمتغير العشوائي X الذي يمثل عدد مرات النجاح من هذا النوع يتبع توزيع ذي الحدين الذي يعطى بالمعادلة التالية:

$$Z=\frac{n1-xn}{\sqrt{nxy}}$$

حيث:

عدد الاستجابات الموجبة: N1

الوسط الحسابي: Xn

الانحراف المعياري: Nxn

احتمال الاستجابات الموجبة: X

احتمال الاستجابات السالبة: Y

مثال1: من اجل التحقق من فرضية أنه لاتوجد فروق دالة احصائيا بين الاطفال في اختياراتهم، تم أخذ 28 طفلا وطبق عليهم اختبار عبارة عن مربعات ومستطيلات. حيث نجح 20 طفلا في التجربة ولم ينجح 08 اطفال.

مثال2: نفرض أن مصنعا لإنتاج الكراريس ينتج 300 كراسا وجدنا 30 كراسا منها غير صالح للاستعمال، وأخذنا عينة من 3 كراريس من انتاج المصنع.

أوجد أحتمال أن يكون من بين الثلاث كراريس المأخوذة كراسين(02) غير صالحين للاستعمال.

مثال2: ألقيت قطعة نقود 4مرات، ماهو احتمال ظهور الصورة 3 مرات.

الحل:

عدد المحاولات n=4

احتمال ظهور الصورة في أي مرة $p=1/2$

احتمال عدم ظهور الصورة في أي مرة $L= 1-p=1/2$

نفرض x هي عدد الصور التي تظهر على السطح العلوي.

اذن $(1/2)x\ (1/2)4-x$

$p(x)= 4kx$

حيث $x= 1,2,3,4$

احتمال ظهور الصورة 3 مرات $p(x=3)$

$=4k3(1/2)3(1/2)1$

$=4/16$

$=1/4$

- توزيع بواسون: يمثل توزيع بواسون أحد التوزيعات الاحتمالية الشائعة في تحليل بيانات المتغيرات المنفصلة. ويختص هذا التوزيع الذي يمثل حالة خاصة لتوزيع ذي الحدين الذي سبق شرحه بالصفات غير المستمرة، وذلك في حالة إذا كان عدد المرات التي يحدث فيها المتغير أو الظاهرة معلوما (فتحي أبوراضي 2000 ص79) وكمثال على ذلك: عدد الهزات الارضية في السنة، عدد حوادث المرور على الطريق السيار(شرق-غرب)، عدد الاخطاء المطبعية في كتاب الجغرافيا للسنة الاولى متوسط...الخ.

  ومن خصائص توزيع بواسون أنه في العادة توزيعا ملتويا التواء موجبان حيث تقابل التكرارات الكبيرة لحدوث المتغير العدد الصغير على مقياس مرات الحدوث، بينما تقابل التكرارات الصغيرة العدد الكبير من مرات الحدوث على نفس المقياس.

- توزيع ستودانت t: إذا كان التوزيع الطبيعي /الاعتدالي يناسب فقط العينات الكبيرة، فإن توزيع t- « student » - الذي يعود إلى مكتشفهGosset الذي كان ينشر أبحاثه باسم مستعار كطالب- صالح للعينات الصغيرة والكبيرة معا، إلا أن استعماله في حالة العينات الصغيرة التي يقل عدد الملاحظات فيها عن 30 يعتبر أكثر فعالية.

توزيع كاي تربيع Distribution khi carré: في البحوث في علم النفس، علم الاجتماع، علوم التربية، وحتى علوم الاعلام والاتصال، يجد الباحث نفسه أمام بيانات كيفية (من مستوى قياس اسمي أو ترتيبي) يتطلب تحليلها في صورة تكرارات مثل: استبيان يحتوي على مجموعة بنود تكون بدائل الاجابة عليها (البنود)، (نعم، لا) أو (موافق، محايد، معارض) أو (راض تماما، راض، غير راض، غير راض تماما)، في هذه الحالة وللكشف والمقارنة بين الفروق المشاهدة وتكرار متوقع الحصول عليه لنفس الاستجابات في المجتمع الأصلي، نطبق توزيع نظري يسمى توزيع كا2 . ويكون الشكل العام لتوزيع (x2) توزيعا ملتويا دوما.

الفرضيات

- الفرضية الإحصائية: هي جملة علمية تعبر عن توقع أو احتمال أو اجابة مؤقتة لسؤال يضعة الباحث ويحاول التحقق منه احصائيا، وهي مرتبطة مباشرة بفرضيات البحث.

تصاغ الفروض الاحصائية في شكل صفري أو بديل.

*الفرض الصفريH0: يفترض الباحث أن العلاقة بين متغيرين أو الفرق بينهما يساوي صفر

مثال:

لا توجد علاقة ارتباطية ذات دلالة احصائية بين متغيري الدافعية والأداء لدى عمال شركة سونلغاز .

$X1-x2=0$

- لا توجد فروق ذات دلالة احصائية في الاداء بين العمال الذين التحقوا بتكوين والعمال الذين لم يلتحقوا بتكوين. $n1-n2=0$

- الفرض البديلH1:يفترض الباحث أن هناك علاقة بين متغيرين أو فروق متوقعة بينهم.

مثال:

- توجد علاقة ارتباطية ذات دلالة احصائية بين متغيري الدافعية والأداء لدى عمال شركة سونلغاز. $X1-x2\neq0$

- توجد فروق ذات دلالة احصائية في الاداء بين العمال الذين التحقوا بتكوين والعمال الذين لم يلتحقوا بتكوين. $n1-n2\neq0$

- مستوى الدلالة: هو المستوى الذي يطمئن عنده الباحث من صحة نتائجه وأنها لا تعود للصدفة. ويتم الكشف عنها من خلال جداول احصائية خاصة وذلك بعد تحديد القيمة المحسوبة. وتكون هذه الجداول غالبا في ملاحق كتب الاحصاء.

وهناك ثلاث مستويات دلالة مقبولة احصائيا:

مستوى دلالة 0.001، أي أن هناك ثقة في النتائج التي توصلت إليها كباحث بنسبة0.999 مقابل شك بنسبة 0.001 أي أن كل 1000مرة نقوم فيها بحساب معامل الارتباط مثلا، هناك 999 مرة صواب مقابل مرة واحدة محتملة للخطأ.

المحاضرة الرابعة: نظرية الاحتمالات:

تعاريف و مفاهيم أساسية

1- التجربة العشوائية (الاحتمالية): هي كل تجربة يمكن تكرارها عددا من المرات، و تكون نتائجها غير محددة سلفا لكونها تعتمد على الصدفة. إذ يمكن معرفة النتائج لكن يصعب تحديد أي منها ستتحقق فعلا.

2- فراغ أو فضاء العينة: و هي مجموعة العناصر الممكنة لنتائج التجربة العشوائية و يرمز لها بالرمز S

مثال: إذا كانت E تجربة إلقاء زهرة النرد مرة واحدة و مراقبة الرقم الذي يظهر على سطحها العلوي بعد استقرارها، فإن هذا الرقم قد يختلف من رمية إلى أخرى، و لكننا نستطيع التأكيد على أنه سينتمي في كل مرة إلى المجموعة التالية: { 1,2,3,4,5,6 } =S

أما إذا كانت E تجربة رمي قطعة نقد مرة واحدة، فإن فراغها أي فضاء العينة هو: { P,F } =S

حيث تمثل P الكتابة و F الصورة

3- الحدث و أنواعه: الحدث هو مجموعة جزئية من فضاء العينة فكما رأينا في المثال السابق يعتبر الرقم 1 حدث و الرقم 2 حدث.....إلخ نفس الشيء بالنسبة للمثال الثاني F هو حدث و أيضا P هو حدث

3-1- الحدث البسيط: و هو الحدث الغير قابل للتجزئة أي هو بمثابة مجموعة جزئية من فراغ إمكانات التجربة، فظهور العدد 3 على وجه زهرة النرد هو بمثابة حدث بسيط من S إذ أنه غير قابل للتجزئة، و حيث أن: { 1,2,3,4,5,6 } =S

3-2- الحدث المركب: و هو الحدث الذي يمكن تفكيكه إلى حوادث أبسط، فعلى سبيل المثال عند رمي قطعة نقد مرتين فإن فضاء العينة $S= \{(P,F), (P,P),(F,F) \}$

3-3- الحدث الأكيد: و هو الحدث الذي يقع دائما عند إجراء التجربة العشوائية، مثل ظهور عدد أقل من 7 في حادثة رمي زهرة النرد.

3-4- الحدث المستحيل: و هو الحدث الغير قابل للتحقق أبدا نتيجة التجربة، فهو المجموعة الخالية المسندة إلى المجموعة الكلية S

مثال: في تجربة إلقاء زهرة النرد، إن حدث ظهور أحد الأرقام الستة لهذه القطعة هو حدث أكيد، في حين أن حدث الحصول على الرقم 8 هو حدث مستحيل، لأنه لا يمكن أن يقع بنتيجة إلقاء زهرة النرد.

3-5- الحدث المتمم: نقول أنه لكل حدث A مرتبط بالتجربة E متمم يتكون من مجموعة الإمكانات غير المحققة ل A و المرتبطة ب E، هذا الحدث الجديد نرمز له ب C أي:

$$CA= \{ ei \in S ; ei \notin A \}$$

واعتمادا على مفهوم فرق حدثين نجد أن: $CA= S - A$

فمثلا في تجربة إلقاء زهرة النرد، إذا كان الحدث $A= \{ 2, 4, 6 \}$

أي حدث حصولنا على رقم زوجي، فإن الحدث المتمم ل: A أي CA يكون: $CA= \{ 1, 3, 5 \}$

3-5- الأحداث المتنافية: بمعنى حدثين أحدهما ينفي الآخر، أي لا يمكن أن يحدثا سوية في نفس الوقت.

فمثلا عند رمي قطعة النقد فإن نتيجة الرمية تكون إما صورة أو كتابة و لا يمكن ظهورهما معا.

3-6- الأحداث المستقلة: و هي الأحداث التي تحدث مستقلة عن الأحداث الأخرى، بمعنى أن كل حدث يحدث بمعزل عن الحدث الآخر.

فمثلا عند رمي قطعة النقد مرتين فإن حادثة ظهور الصورة في الرمية الأولى لا تؤثر على ظهور صورة الكتابة في الرمية الثانية.

القوانين الأساسية في نظرية الاحتمالات

1- القانون الأساسي: إن احتمال وقوع أي حدث عبارة عن قسمة عدد الحالات المحققة للحدث على عدد الحالات الكلية. الاحتمال يتراوح بين الصفر و الواحد

2- جمع الاحتمالات: إن دراسة جمع الاحتمالات يتطلب التساؤل عن العلاقة الموجودة بين الأحداث كونها متنافية أو غير متنافية، و على أساس الإجابة تكون قاعدة الجمع

2-1- الأحداث المتنافية: إذا كان A و B حدثين متنافيين، فإن تحقق إجتماعهما يساوي مجموع احتمالهما أي أن: $P= (A \cup B) = P (A)+ P(B)$

و تعرف هذه الصيغة باسم القاعدة الخاصة للجمع.

مثال: لنفرض أننا وضعنا في صندوق خمس كرات: إثنتان بيضاويتين وواحدة سوداء و إثنتين حمراويتين و سحبنا واحدة. فما هو احتمال الحصول على كرة بيضاء أو سوداء؟

الحل:

إن احتمال الحصول على المطلوب يساوي إلى احتمال الحصول على كل منهما، فاحتمال الحصول على كرة بيضاء يساوي: $P (B) = 2/5$

أما احتمال الحصول على كرة سوداء فيساوي: $P (N) = 1/5$

واحتمال الحصول على أحدهما يساوي: $P (B \cup N) = P (B) + P (N) = 2/5+1/5=3/5$

و يمكننا ببساطة أن نصل إلى التعميم التالي:

إن احتمال تحقق إجماع أحداث متنافية يساوي مجموع إحتمالات تلك الأحداث.

مثال:

لدينا صندوق قطع غيار به 200 قطعة، منها مئة و خمسون من النوع الجيد (الحدث A) و ثلاثون قطعة من النوع المتوسط (الحدث B) و ست عشرة قطعة من النوع الردئ (الحدث C) و أربع قطع عديمة النفع (الحدث D) ، سحبنا قطعة من هذا الصندوق بشكل عشوائي، فما هو إحتمال أن تكون هذه القطعة جيدة أو متوسطة أو رديئة؟

الحل:

لا يمكن أن تكون القطعة المسحوبة جيدة، متوسطة و رديئة في آن واحد، أي أن هذه الحالات لا تحدث مع بعضها، فعند سحب قطعة من النوع الجيد يعني نفي السحب من النوع المتوسط و الردئ، و عند سحب قطعة من النوع الردئ يعني نفي السحب من النوع الجيد والمتوسط، و عليه:

$$P(A) = 150/200 = 0.75$$

$$P(B) = 30/200 = 0.15$$

$$P(C) = 16/200 = 0.08$$

و منه: $P(A \cup B \cup C) = P(A) + P(B) + P(C)$

$$= 0.75 + 0.15 + 0.68 = 0.98$$

2-2- الأحداث غير المتنافية:

إذا كان A و B حدثين غير متنافيين، فإن إحتمال وقوع أحدهما هو عبارة عن حاصل جمع إحتمال وقوع كل منهما مع إبعاد إحتمال وقوعهما معا بآن واحد، أي أن:

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

و تعرف هذه الصيغة باسم القانون العام للجمع

مثال:

يقوم شخص بإلقاء زهرة النرد، فما هو احتمال الحصول على رقم يكون فردي أو يقبل القسمة على العدد 3.

نلاحظ هنا أن حصولنا على رقم فردي لا يمنع من الحصول على رقم يقبل القسمة على العدد 3، و في نفس الوقت، حيث نجد أن:

- عدد النتائج االممكنة، التي تمثل العدد الفردي (الحدث A): { 1, 3, 5 } =A

- عدد النتائج الممكنة و التي تمثل عدد يقبل القسمة على 3 (الحدث B): { 3, 6 } =B

إن العدد الفردي و الذي يقبل القسمة على 3 هو العدد 3 فقط، و على ذلك فإن:

$$P(A)= 3/6 \qquad P(B)= 2/6$$

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

$$= 3/6+ 2/6 - 1/6= 4/6$$

ثالثا- ضرب الاحتمالات:

3-1- قانون الضرب للحوادث المستقلة: إناحتمال وقوع حدثين مستقلين معا في آن واحد يساوي إلى حاصل ضرب احتمال وقوع كل منهما و نعبر عن ذلك بالعلاقة التالية: $P(A\cap B)= P(A) \times P(B)$

مثال:

صندوق يحوي 10 مصابيح كهربائية منها 7 صالحة للإضاءة و 3 غير صالحة للإضاءة، قام أحد الأشخاص بسحب مصباحين عشوائيا من الصندوق مع إرجاع المصباح الأول إلى الصندوق لنجد مايلي:

أ- إحتمال كونهما صالحين للإضاءة

ب- إحتمال كونهما غير صالحين للإضاءة

ج- إحتمال كون الأول صالح للإضاءة و الثاني غير صالح

الحل:

أ- احتمال كون المصباحين صالحين للإضاءة:

- الأول صالح للإضاءة: الحدث A

- الثاني صالح للإضاءة: الحدث B

- الأول صالح للإضاءة و الثاني صالح للإضاءة: الحدث $A \cap B$

و عليه: $P(A \cap B) = P(A) \times P(B)$

$P(A) = 7/10$ $P(B) = 7/10$

$P(A \cap B) = 7/10 \times 7/10 = 49/100 = 0.49$

ب- احتمال كون المصباحين غير صالحين للإضاءة

- الأول غير صالح للإضاءة: الحدث C

- الثاني غير صالح للإضاءة: الحدث D

- الأول غير صالح للإضاءة و الثاني غير صالح للإضاءة: $C \cap D$

$P(C \cap D) = P(C) \times P(D)$

$P(C) = 3/10$ $P(D) = 3/10$

$P(C \cap D) = 3/10 \times 3/10$

ج- إحتمال كون المصباح الأول صالح و الثاني غير صالح

$P(A \cap C) = P(A) \times P(C)$

$= 7/10 \times 3/10 = 21/100$

يمكننا أن نعمم مفهوم الاستقلال هذا على أكثر من حدثين شريطة أن يكون استقلالها محقق مثنى مثنى.

3-2- قانون الضرب للحوادث غير المستقلة: إذا كان لدينا الحدثان A و B و كان وقوع الحدث B مشروطا بوقوع الحدث A فإن احتمال وقوعهما معا يساوي جداء احتمال وقوع الأول باحتمال وقوع الثاني بعد حدوث الأول.

و يمكن التعبير عن ذلك بما يلي: يعتبر الحدثان غير مستقلين إذا كان وقوع أحدهما مرتبطا بطريقة ما بوقوع الآخر: $P(A\cap B)= P(A) \times P(B/A)$

أو $P(A\cap B)= P(B) \times P(A / B)$

و ذلك حسب الحدث المشروط بوقوع الحدث الآخر، و تسمى هذه العلاقة عادة بقاعدة الضرب.

فإذا كان $P(B)\neq 0$، فإن الاحتمال الشرطي لوقوع الحدث A بعد معرفة وقوع الحدث B يعطى بالعلاقة التالية: $P(A / B)= P(A\cap B) /P(B)$

و بشكل مشابه يحدد الاحتمال الشرطي لوقوع الحدث B بافتراض وقوع الحدث A، علما أن $P(A) \neq 0$ بالعلاقة التالية: $P(B/A)= P(A\cap B) / P(A)$

مثال:

قام أحد الأفراد بسحب قطعتين من إنتاج مصنع ما، بصورة عشوائية، من داخل صندوق به 100 وحدة جيدة و 10 غير جيدة، فإذا كان السحب تعاقبيا و دون إرجاع القطع المسحوبة. و لنحسب ما يلي:

أ- احتمال كون القطعتين المسحوبتين جيدتين

ب- احتمال كون القطعتين المسحوبتين غير جيدتين

ج- احتمال كون إحداهما جيدة و الأخرى غير جيدة

الحل:

نظرا لكون السحب يتم على التوالي (بالتعاقب) دون إرجاع القطع المسحوبة فهذا يعني أن الأحداث غير مستقلة واحتمال وقوعها يتضمن احتمالا شرطيا و عدم إرجاع القطعة المسحوبة الأولى يعني إنقاص (تخفيض) عدد القطع الكلية و النوعية (جيدة– غير جيدة) بوحدة واحدة و على ذلك فإن:

أ– احتمال كون القطعتين المسحوبتين جيدتين:

– الأولى جيدة A

– الثانية جيدة B

– الأولى جيدة و الثانية جيدة $A \cap B$

و عليه: $P(A \cap B) = P(A) \times P(B/A)$

$P(A)=100/110 \qquad P(B/A) = 99/109 \qquad P(A \cap B) = (100/110) \times (99/109) = 0.83$

ب– احتمال كون الوحدتين المسحوبتين غير جيدتين

– الأولى غير جيدة C

– الثانية غير جيدة D

– الأولى غير جيدة و الثانية غير جيدة $C \cap D$

و عليه: $P(C \cap D) = P(C) \times P(D/C)$

$P(C) = 10/110 \qquad P(D) = 9/109$

$P(C \cap D) = (10/110) \times (9/109) = 0.0075$

ج– احتمال كون أحدهما جيدة و الأخرى غير جيدة

هنا العبارة وردت دون تحديد للموقع بالنسبة للجودة (إحداهما) و بالتالي نكون بصدد وضعيتين:

$$P(C \cap \overline{D}) = (10/110) \times (10/109) = 0.083$$

و بالتالي فاحتمال كون إحداهما جيدة و الأخرى غير جيدة هو عبارة عن مجموع الوضعيتين، مع العلم أنهما متنافيتان:

$$P(A \cap \overline{C}) \cup (C \cap \overline{A}) = P(A \cap \overline{C}) + (C \cap \overline{A}) = 0.083 + 0.083 = 0.166$$