

علم الاحصاء: هو ذلك العلم الذي يعمل على استخدام الاسلوب العلمي في طرق جمع البيانات وتبويبها وتلخيصها وعرضها وتحليلها بهدف الوصول منها على استنتاجات وقرارات مناسبة

مفاهيم الاحصاء :

1-البيانات:

أ-البيانات غيرالمبوبة (Ungrouped data):

هي البيانات الاولية (Raw data) التي جمعت ولم تبوب

ب-البيانات المبوبة (grouped data):

هي البيانات التي نظمت في جدول توزيع تكراري.

2-المتغير :هي أي ظاهرة توجد اختلافات بين مفرداتها ويرمز لها بالرمز الخ والمتغيرات نوعان

أ-المتغيرات الوصفية (نوعية):

وهي المتغيرات التي يمكن قياسها بوسائل قياس مألوفة او هي تلك الظواهر والصفات التي لايمكن قياسها مباشرة بالارقام العددية مثل الحالة الاجتماعية (غني،فقير)

ب-متغيرات كمية:

هي تلك الظواهر او الصفات التي يمكن قياسها مباشرة بالارقام العددية مثل صفة الطول والوزن والعمروتقسم المتغيرات الكمية الى:

1-متغيرات المستمرة :وهي المتغيرات التي تاخذ المفردة فيها أي قيمة رقمية ضمن مدى معين لو افترضنا ان اطوال مجموعة من الاشخاص تتراوح بين 130-170سم

نكتب المتغير بصيغة $130 \leq X \leq 170$

2- متغيرات المتقطعة (المنفصلة) :وهي المتغيرات التي تاخذ المفردة فيها قيما

متباعدة لو افترضنا ان عدد افراد الاسرة في اربعة عوائل هي 5,4,3,2

نكتب المتغير بصيغة $X = 2, 3, 4, 5$

المجتمع :عبارة عن جميع القيم او المفردات التي يمكن ان ياخذها المتغير والمجتمع
اما ان يكون :

أ-مجتمع محدد: أي يمكن حصر عدد مفرداته مثلا عدد وحدات منتجة في مصنع معين.

ب- مجتمع غير محدد: وهو المجتمع الذي يصعب حصر عدد مفرداته مثلا عدد البكتريا في حقل معين.

4- العينة :عبارة عن مجموعة من المفردات يتم اختيارها بطريقة معينة من المجتمع

المجموع (Summation) :

اذا كانت (x_1, x_2, \dots, x_n) تمثل مفردات المتغير X فان مجموع هذه
المفردات يعبر عنها بالرمز : $\sum X_i$

اي ان : $\sum X_i = x_1 + x_2 + \dots + x_n$

EX : $X = 6, 7, 4, 3$, find $\sum X_i, \sum X_i^2$

SOL: $\sum X_i = 6 + 7 + 4 + 3 = 20$

$\sum X_i^2 = 36 + 49 + 16 + 9 = 110$

EX: $X = 5, 3, 7, 1$ find $(\sum X_i)^2$

العرض الجدولي (Tabular Presentation):

جدول التوزيع التكراري (Frequency Distribution):

هو جدول بسيط يتكون من عمودين ،العمود الاول تقسم فيه قيم المتغير الى مجموعات تدعى الفئات (Classes) والثاني يبين فيه عدد مفردات كل فئة يسمى التكرار (Frequency) ولكل فئة حدان الحد الاعلى الحقيقي والحد الادنى الحقيقي فاذا كانت حدود الفئات اعدادا صحيحة فان الحد الادنى الحقيقي لاي فئة يكون

مساويا للحد الأدنى لتلك الفئة مطروحا منه (0.5) والحد الأعلى الحقيقي لأي فئة يكون مساويا للحد الأعلى مضافا إليه (0.5)

1-جدول البسيط: وهو الجدول الذي توزع فيه البيانات حسب صفة واحدة ويتألف عادة من عمودين الأول يمثل تقسيمات الصفة أو الظاهرة إلى فئات أو مجموعات والثاني يبين عدد المفردات التابعة لكل فئة أو مجموعة ،

مثال/ الجدول التالي يمثل توزيع عدد من طلبة جامعة ما حسب أوزانهم:

الفئات	Fi
60-62	85
63-65	15
66-68	45
69-71	27
72-74	8
المجموع	180

2-جدول المزدوج :

وهو الجدول الذي توزع فيه البيانات حسب صفتين أو ظاهرتين أو أكثر في نفس الوقت مثلا الجدول الزوجي (لصفتين) يتألف من الصفوف تمثل فئات أو مجاميع إحدى الصفتين و الأعمدة تمثل فئات أو مجاميع الصفة الأخرى أما المربعات التي تقابل الصفوف والأعمدة تحتوي على عدد المفردات أو التكرارات المشتركة في فئات ومجاميع كلا الصفتين،

مثال/ الجدول الآتي يمثل توزيع عدد من طلبة كلية ما حسب صفتي الطول والوزن

المجموع	71-80	61-70	51-60	
30	4	6	20	121-140
52	10	40	2	141-160
18	10	6	2	161-18
100	24	52	24	المجموع

طول الفئة : (Class Length)

هو مقدار المدى بين حدي الفئة وهي عبارة عن الفرق بين الحدين الأعلى أو الحدين الأدنى للفئتين المتتاليتين.

مركز الفئة (Class mid -Point) : وهو عبارة عن منتصف المدى بين الحدي
الفئة اي :

$$\text{مركز الفئة} = (\text{الحد الاعلى} + \text{الحد الادنى}) / 2$$

التكرار النسبي :وهي تبين الاهمية النسبة لكل فئة ويحسب كالاتي :

$$\text{التكرار النسبي لاي فئة} = (\text{تكرار الفئة} / \text{مجموع التكرارات}) = f_i / \sum f_i$$

حيث تشير f_i الى التكرار الفئة و $\sum f_i$ الى مجموع التكرارات الكلية

$$\text{التكرار المئوي} = (\text{التكرار النسبي} \times 100)$$

جدول التوزيع التكراري المتجمع الصاعد (Less than Clumulative -dist)

هو الجدول الذي يعطينا عدد المفردات التي تقل قيمتها عن الحد الادنى لفئة معينة.

المدرج التكراري (Histogram)

عبارة عن مستطيلات رأسية تمتد قواعدها على المحور الافقي لتمثل اطوال الفئات
بينما ارتفاعاتها تمثل تكرارات الفئات ولرسم مدرج تكراري تتبع الخطوات الاتية :

1- نرسم المحور الافقي والمحور العمودي

2- يتم تقسيم المحور الافقي الى اقسام متساوي بمقياس رسم مناسب بحيث

يشمل جميع الحدود الحقيقية للفئات ويفضل ترك مسافة صغيرة بين نقطة

الاصل والحد الادنى للفئة الاولى ويقسم المحور العمودي الى اقسام متساوية

بحيث تشمل اكبر التكرارات

3- يرسم على كل فئة مستطيل راسي تمثل قاعدته طول تلك الفئة وارتفاعه

يمثل تكرار تلك الفئة .

المضلع التكراري: (Freq . Polygon)

عبارة عن خطوط مستقيمة متكسرة تصل بين نقاط كل نقطة فيها واقعة فوق مركز

فئة على ارتفاع يمثل تكرار تلك الفئة وعادة يقفل المضلع بايصال بدايته بالمحور

الافقي بمركز فئة خيالية واقعة الى يسار اول فئة تكرارها صفر ونصل نهاية

المضلع بالمحور الافقي بمركز فئة خالية واقعة الى يمين اخر فئة وتكرارها صفر

ايضا.

ملاحظة :لرسم المضلع التكراري يجب رسم المرج التكراري اولا ثم تصنف القواعد العليا للمستطيلات والتي تمثل مراكز الفئات بنقاط ثم نصل هذه النقاط بمستقيمات.

المنحني التكراري (Freg Carve):

لايختلف فكرة رسم المنحني التكراري(Freg Carve) عن المضلع التكراري من حيث الاسلوب لكل الفرق الوحيد ما بينهما هو انه بدلا من توصيل النقاط (مركز الفئة ،التكرار) بمستقيمات فانه يتم تمرير منحني ما بين هذه النقاط هذا المنحني يمثل المنحني التكراري.

التمثيل البياني لجدول التوزيع التكراري التجميعي :

يستخدم لذلك المضلع التكراري التجميعي وهو عبارة عن خطوط مستقيمة متكسرة تصل بين نقاط واقعة فوق الحدود الحقيقية للفئات واعلى ارتفاع يمثل التكرار التجميعي وهنالك نوعان من المضلع التكراري التجميعي :

1- المضلع التكراري التجميعي التصاعدي:

لرسمه نتبع الخطوات الاتية :

أ-يتدرج المحور الافقي الى اقسام متساوية تمثل جميع حدود الفئات ويقسم المحور العمودي الى اقسام متساوية بحيث تشمل على اكبر التكرارات التجميعية وهي المجموع الكلي للتكرارات.

ب- وضع نقطة امام كل حد فئة ويعادل ارتفاعها التكرار التجميعي التصاعدي لذلك الحد ثم نوصل تلك النقاط بخطوط مستقيمة أي ان المضلع يبدأ من الصفر وينتهي باعلى نقطة .

2-المضلع التكراري التجميعي التنازلي:

ويرسم بنفس الطريقة السابقة غير ان هذا المضلع يبدأ من اعلى نقطة أي(مجموع التكرارات) وينتهي بالصفر.

مثال/ الجدول التالي يمثل توزيع عدد من طلبة جامعة ما حسب اوزانهم:

الفئات	Fi
60-62	5
63-65	18
66-68	42
69-71	27
72-74	8
المجموع	100

اوجد: 1- الحدود الحقيقية للفئات

2- مراكز الفئات

3- التكرار النسبي والمئوي

4- جدول التوزيع التكراري التجميعي الصاعد مع الرسم

5- جدول التوزيع التكراري التجميعي النازل مع الرسم

6- ارسم المدرج التكراري والمضلع التكراري

الحل:

الفئات	تكرار	Xi	الحدود الحقيقية	التكرار النسبي	التكرار المئوي	التكرار التجميعي الصاعد	التكرار التجميعي النازل
60-62	5	61.5	59.5 - 62.5	0.05	5	0	100
63-65	18	64	62.5- 65.5	0.18	18	5	95
66-68	42	67	65.5- 68.5	0.42	42	23	77
69-71	27	70	68.5- 71.5	0.27	27	65	35
72-74	8	73	71.5-	0.08	8	92	8

			74.5				
المجموع	100					100	0

ثم نعين النقاط على محور و ونرسم المنحني الصاعد والمنحني النازل

الفصل الثاني

1- مقاييس النزعة المركزية

1-الوسط الحسابي :وهي القيمة الناتجة من قسمة مجموع تلك القيم على عددها

أ-إذا كانت البيانات غير مبوبة:إذا كان لدينا n من القيم أي (x_1, x_2, \dots, x_n) فان الوسط الحسابي لها هو $X = \sum x_i / n$

ب-البيانات المبوبة: إذا كان لدينا (x_1, x_2, \dots, x_n) تمثل مراكز الفئات في جدول التوزيع التكراري مع تكراراتها (f_1, f_2, \dots, f_n) على التوالي فان الوسط الحسابي لها هو $X = \sum f_i x_i / \sum f_i$

مثال: من جدول التوزيع التكراري الاتي جد الوسط الحسابي للبيانات التالية:

الفئات	F_i	X_i	$F_i X_i$
31-40	1	35.5	35.5
41-50	2	45.5	91.0
51-60	5	55.5	277.5
61-70	15	65.5	982.5
71-80	25	75.5	1887.5
81-90	20	85.5	1710.5
91-100	12	95.5	1146.5

$$X_i = \sum f_i x_i / \sum f_i = 6130 / 80 = 76.62$$

2-الوسيط:

أ-إذا كانت البيانات غير مبوبة:إذا كان لدينا n من القيم (x_1, x_2, \dots, x_n) ترتب القيم ترتيبا تصاعديا او تنازليا:1-إذا كان n عددا فرديا فان الوسيط هو القيمة التالية

$$Me = (n+1)/2$$

2-إذا كان n عددا زوجيا فان الوسيط يكون: $(n/2, n/2 + 1)$

مثال 1: جد الوسيط للقيم الاتية " $X=80,82,76,87,84$ "

الحل/نرتب القيم ترتيبا تصاعديا او تنازليا ثم نلاحظ ان عددا فرديا نطبق القانون

$$(n+1)/2=82 \text{ هو الوسيط هو } X=76,80,82,84,87, n=5$$

مثال 2/اذا كانت n زوجية نرتب القيم تصاعدي او تنازلي ثم نطبق القانون التالي

ليكن $n=8$ و $X=5,4,8,7,3,12,9,2$ اذن نطبق القانون التالي $n/2, n/2+1$

$$8/2=4, 4+1=5 \text{ اذن } X=2,3,4,5,7,8,9,12, n=8,$$

$$5+7=12/2=6 \text{ اذن الوسيط هو } 6$$

ب-البيانات مبوبة :اذا كانت (x_1, x_2, \dots, x_n) تمثل مراكز الفئات مع تكراراتها (f_1, f_2, \dots, f_n) على التوالي فان قيمة الوسيط لهذه البيانات هي

$$Me = L1 + (((f_i/2) - F_i) / f_i) \times w$$

حيث ان $(L1)$ يمثل الحد الادنى للفئة الوسيط

و (f_i) تمثل مجموع التكرارات

و (F_i) تمثل التكرار المتجمع الصاعد عند بداية فئة الوسيط

و (f_i) تمثل فئة الوسيط

و (w) تمثل طول فئة الوسيط

مثال/من جدول التوزيع التكراري جد الوسيط للبيانات التالية"

الفئات	F_i	التكرار التجميعي الصاعد
60-62	5	5
63-65	18	23
66-68	42	65
69-71	27	92
72-74	8	100

نجد الوسيط باستخدام القانون التالي

ثم نلاحظ القيمة واقعة بين (23-65) من جدول التوزيع المتجمع الصاعد وبعد ذلك نعين فئة الوسيط من الفئات الاصلية (66-68) ونجد الحدود الحقيقية لها ونجد تكرار الفئة الوسيط ($f_i=42$) وطول الفئة (w) وبعد نطبق القانون

$$Me = L1 + (((\sum f_i / 2) - F_i) / f_i) \times w$$

3- المنوال:

ا- البيانات غير مبوبة: اذا كان لدينا n من القيم (x_1, x_2, \dots, x_n) فان المنوال لها يكون الاكثر تكرارا ويرمز له ب M

مثال/ اذا كانت $X = 3, 5, 2, 6, 5, 9, 2, 5, 8, 6$ جد المنوال لها؟

الحل/ المنوال هو $= 5$ لانه الاكثر تكرارا

ب- البيانات المبوبة: نستخدم هذه الطريقة عندما تكون الفئات مستمرة أي اذا كانت لدينا (x_1, x_2, \dots, x_n) تمثل مراكز الفئات في جدول توزيع تكراري مع تكراراتها (f_1, f_2, \dots, f_n) على التوالي فان المنوال هو

$$M = L1 + (d1 / d1 + d2) * w$$

حيث ان فئة المنوال هي التي تمثل اكبر التكرارات

وان ($L1$) تمثل الحد الادنى الحقيقي لفئة المنوال

و ($d1$) هو الفرق بين تكرار الفئة المنوال والفئة السابقة لها

و ($d2$) هو الفرق بين تكرار الفئة المنوال والفئة اللاحقة لها

و (w) هو طول الفئة

اما اذا كانت البيانات متقطعة نستخدم القانون التالي المنوال = (الحد الادنى + الحد الاعلى) / 2

مثال/ من الجدول السابق جد المنوال له"

$$M = 66 + 68 / 2 = 67$$

في حالة المتقطعة نستخدم القانون التالي المنوال = (الحد الادنى + الحد الاعلى) / 2

بمعنى اخر المنوال يمثل مركز الفئة التي تقابل اكبر التكرارات

مثال/من الجدول التالي جد المنوال له .

الفئات	20-	30-	40-	50-	60-	70-	80 - 90
F_i	2	8	10	20	13	12	5

الحل/نطبق القانون $M=L1+(d1/d1+d2) \times w$

$$M= 49.5+(10/10+7) 10=49.5+5.88=55.38$$

2-مقاييس التشتت او الاختلاف:

ونقصد به التباعد او التقارب الموجود بين قيم المشاهدات لمتغير ما وايضا هي مقياس مدى تشتت قيم المشاهدات عن وسطها فكلما كان التشتت كبيرا دل ذلك على

عدم تجانس بين القيم والعكس صحيح وهناك عدة مقاييس للتشتت اهمها:

1-مقاييس التشتت المطلق: أي التي تكون وحداتها نفس وحدات القيم الاصلية واهمها

ا-المدى (The Rang)

ب-الانحراف المتوسط)

ج-التباين والانحراف المعياري(القياسي)

2-مقاييس التشتت النسبي:وهو الذي يكون خالي من وحدات القياس واهمها:

"معامل الاختلاف"

المدى: لمجموعة من القيم هو الفرق بين اعلى قيمة واقل قيمة من تلك المجموعة

ويرمز لها ب(R)

مثال/لتكن $X=9,4,7,8,16, 10,20,2$

الحل/ $R= 20-2=18$

هذه في حالة اذا كانت البيانات غير مبوبة نستخدم هذه الطريقة اما اذا كانت البيانات

مبوبة نستخدم الجدول ونجد مراكز الفئات ثم نجد الفرق بين اعلى قيمة واقل قيمة من مراكز الفئات يمثل المدى

الانحراف المتوسط: اذا كانت البيانات غير مبوبة نستخدم القانون التالي:

$$M_D = \frac{\sum |x_i - \bar{x}|}{N} \text{ وحيث } M_D \text{ تمثل الانحراف المتوسط}$$

والسبب في اخذ الانحراف المطلق هو ان ابقاء الاشارات تكون موجبة لان السالب

تجعل مجموع الانحرافات يساوي صفرا

مثال/ لتكن $X = 9, 8, 6, 5, 7$ جد الانحراف المتوسط لهذه القيم؟

$$\bar{x} = \frac{\sum x_i}{n} = \frac{9+8+6+5+7}{5} = \frac{35}{5} = 7$$

$$M_D = \frac{\sum |x_i - \bar{x}|}{n}$$

$$M_D = \frac{2+1+1+2+0}{5} = \frac{6}{5}$$

X_i	$x_i - \bar{x}$	$ x_i - \bar{x} $
9	2	2
8	1	1
6	-1	1
5	-2	2
7	0	0

اما اذا كانت البيانات مبوبة نطبق القانون التالي:

$$M_D = \frac{\sum f_i |x_i - \bar{x}|}{\sum f_i}$$

مثال/ من جدول البيانات التالي جد الانحراف المتوسط

الفئات	F_i	X_i	$F_i x_i$	$ x_i - \bar{x} $	$F_i x_i - \bar{x} $
10-18	5	14	70	26.5	132.5
19-27	8	23	184	17.5	140.0
28-36	10	32	320	8.5	85.0
37-45	20	41	820	0.5	10.0
46-54	17	50	850	9.5	161.5
55-63	10	59	590	18.5	185.0

$$x = \sum f_i x_i / \sum f_i = 2834 / 70 = 40.48 \text{ نجد الوسط الحسابي}$$

ثانيًا نطبق القانون :

$$M_D = \sum f_i |x_i - x| / \sum f_i = 714.0 / 70 = 10.2$$

2- التباين او الانحراف المعياري: في حالة البيانات غير مبوبة

اذا كان لدينا (n) من المشاهدات (x_1, x_2, \dots, x_n) فان التباين له والذي

$$S^2 = \sum (x_i - x)^2 / n - 1 \text{ يكون له } (S^2) \text{ نرمز له}$$

ويمكن ايجاد التباين بطريقة مختصرة وهي

$$S^2 = [\sum x_i^2 - (\sum x_i / n)^2] / n - 1$$

وان الجذر التربيعي للتباين يسمى بالانحراف المعياري ويرمز له

ب(S)

مثال/جد التباين والانحراف المعياري للقيم التالية: $x=4,5,7,9,3$

$$x_i^2 = 16, 25, 49, 81, 9 \text{ اولاً نجد}$$

$$\sum x_i^2 = 16 + 25 + 49 + 81 + 9 = 432$$

$$\sum x_i = 4 + 5 + 7 + 9 + 3 = 28 \text{ ثم نجد مجموع قيم } x_i$$

$$S^2 = 432 - (28)^2 / 5 - 1 = 432 - 4 = 428 \text{ ونطبق القانون:}$$

$$S = \text{الانحراف المعياري}$$

اما بالنسبة للبيانات المبوبة نستخدم القانون التالي:

$$S^2 = \sum (f_i x_i^2 - \sum (f_i x_i)^2 / f_i) / \sum f_i - 1$$

$$S^2 = \sum (f_i |x_i - x|^2) / \sum f_i - 1 \text{ او}$$

مثال البيانات المبوبة/جد التباين والانحراف المعياري للبيانات

التالية:

الفئات	F_i	X_i	Fix_i	X_i^2	$F_i X_i$
60-62	5	61	305	3721	18605
63-65	18	64	1152	4096	73729
66-68	42	67	2814	4489	188538
69-71	27	70	1890	4900	132300
72-74	8	73	584	5329	42632

$$S^2 = \frac{\sum fix_i^2 - \frac{(\sum fix_i)^2}{\sum f_i}}{\sum f_i - 1} =$$

$$S^2 = 455803 - \frac{(6745)^2}{100/99} = 8.6$$

اما الانحراف المعياري فهو $S = 2.9$

4-مقاييس اخرى للتشتت المطلق :هناك بعض المقاييس لاخرى لقياس التشتت ولكنها قليلة الاستعمال وتسمى تشبيهات المدى

(_ Quasi- rang es) مثل :

أ-نصف المدى الربيعي(الانحراف الربيعي)

(Semi _ int erquartile rang e)

ب-نصف المدى العشري او المئيني.

مقاييس التشتت النسبي :ان مقاييس التشتت النسبي لها اهميتها عند مقارنة التشتت لمجموعتين او اكثر تختلف في وحدات القياس لقيمتها لان مقاييس التشتت تكون خالية من وحدات القياس واهم هذه المقاييس هي :

1-معامل الاختلاف *Coefficient of Variation* :اذا كان لدينا

(X, S) هما الانحراف المعياري والوسط الحسابي لمجموعة من القيم

على التوالي فان معامل الاختلاف لها هو ويرمز له ب $C.V$

$$C.V = S/X * 100$$

مثال/من نتائج الامتحانات للمادتين الرياضيات والفيزياء حصلنا على

المعلومات الاتية مادة الرياضيات كان الوسط الحسابي

له(78)والانحراف المعياري هو(8) والوسط الحسابي لفيزياء

هو(73)والانحراف هو(7.6) في أي المادتين كان التشتت اكثر؟

الحل/نطبق القانون التالي $C.V = S/X*100$

لمادة الرياضيات $C.V = 8/78*100 = 10.25 \%$

لمادة الفيزياء $C.V = 7.6/ 73 *100 = 10.41 \%$

أي ان تشتت للمادة الفيزياء اكثر

مثال/من نتائج الامتحانات للمواد التالية احصاء وكيمياء و فيزياء
حصلنا على المعلومات التالية وسط الحسابي لمادة الاحصاء
هي (81) والانحراف هو (9.2) لمادة الكيمياء الوسط الحسابي
هو (76) وانحراف (8.12) ولمادة الفيزياء هو (73) وانحراف (7.6)
في أي من هذه المواد كان تشتت الدرجات اكثر؟

الحل/ نطبق القانون $C.V = S/x*100$

لاحصاء $C.V = 9.2/81*100 = 11.35$

للكيمياء $C.V = 8.12/76*100 = 10.68$

للفيزياء $C.V = 7.6/73*100 = 10.41$

نلاحظ ان تشتت للمادة الاحصاء اكثر من باقي المواد

الدرجة القياسية :

في كثير من الاحيان نحتاج الى المقارنة مفردتين من مجموعتين
مختلفتين وفي هذه الحالة يجب تحويل وحدات كل مفردة الى وحدات
قياسية حتى تكون المقارنة ذات معنى وذلك باستخدام الوسط الحسابي
او الانحراف المعياري لكل مجموعة وتسمى القيمة (Z_i) درجة قياسية
اذا كانت .

$$Z_i = X_i - x / s$$

مثال / حصل طالب على درجة (84) في الامتحان النهائي
بالرياضيات علما بان الوسط الحسابي في امتحان الرياضيات لجميع
الطلبة كان (76) وانحراف قياسي قدره (10) اما في امتحان الكيمياء
فقد حصل الطالب على درجة (90) حيث كان الوسط الحسابي لجميع

الطالبة (82) وبانحراف(16) ففي أي الموضوعين كانت قابلة هذا الطالب اعلى؟

الحل/ نحول هاتين الدرجتين الى درجات قياسية با استخدام القانون

$$Z_i = \frac{X_i - \bar{x}}{s}$$

$$Z_i = \frac{84 - 76}{10} = 0.8 \quad \text{للرياضيات}$$

$$Z_i = \frac{90 - 82}{16} = 0.5 \quad \text{للفيزياء}$$

ومن هذا يتضح بان قابلية الطالب في الرياضيات اعلى من الفيزياء على الرغم من ان درجته في الفيزياء اعلى

الفصل الثالث

مبادئ نظرية الاحتمال

مفاهيم اساسية: 1- التجربة العشوائية :

هي التجربة التي لا يمكن معرفة نتائجها لخضوعها لقوانين الاحتمال

2- فضاء العينة : هو مجموعة كل النتائج الممكنة لتجربة تحدث عن طريق الصدفة ويرمز لها بالرمز (S)

مثال/ عند رمي قطعة نقود مرة واحدة فان فضاء العينة يتكون من نتيجتين ممكنتين هما $S = [H, T]$ حيث يرمز H للصورة، T للكتابة

اما اذارمينا قطعتين من النقود فان فضاء العينة سيكون من اربعة نتائج ممكنة هي $S = [HH, HT, TH, TT]$

الحادث: هو مجموعة جزئية من فضاء العينة ويرمز له ب (Ei)

مثال/ لرمي قطعة نقود مرة واحدة $S = [H, T]$ فان

$$E_1 = [H], E_2 = [T], E_3 = [H, T]$$

1- الحوادث المتنافية :

يقال عن الحادثتين $E1, E2$ انهما متنافيان اذا استحال حدوثهما معا
فمثلا : عند رمي قطعة نقود فمن المستحيل الحصول على صورة
وكتابة في نفس الوقت

2-الحوادث المستقلة:

هي الحوادث التي اذا وقع احدهما لا يؤثر في وقوع الاخر فمثلا:

عند رمي قطعتين نقود فان نتيجة القطعة الاولى لا تؤثر بنتيجة
القطعة الثانية.

3الحوادث غير المستقلة:

هي الحوادث التي اذا وقع احدهما يؤثر في وقوع الحوادث الاخرى فمثلا:في
حالة صندوق به كرات بيضاء وسوداء فعند سحب كرتان على التوالي بحيث لا
تعاد الكرة الاولى فان نتيجة السحبة الثانية تتاثر بنتيجة السحبة الاولى لذا
فالحادثتان غير مستقلتين أي معتمدة

مضروب n : Factorial n

مضروب n ويرمز له ب $(n!)$ ويعرف بانه

$$n! = n (n-1) (n-2)(n-3) \dots \dots \dots 1$$

ملاحظة: $0! = 1, 1! = 1, n! = n (n-1)$

مثال/جد قيمة $5!$!

$$5! = 5(4)(3)(2)(1) = 120 \quad \text{الحل/}$$

التباديل:

هو ترتيب n من العناصر في شكل معين ويرمز له ب (Pr^n) أي تبادل من

$$Pr = n! / (n-r)!$$

مثال/اذا كان لدينا اربعة حروف هي $[A,B,C,D]$ واختير منها حرفان فما هي عدد
الطرق التي يمكن بها اختيار هذين الحرفين

$$Pr^n = 4! / (4-2)! = 4(3)(2)(1) / 2(1) = 24 / 2 = 12 \quad \text{الحل/}$$

أي ان عدد الطرق = 12 وهذه الطرق هي
[AB,AC,AD,BC,BD,CD,BA,CA,DA,CB,DB,DC]

حيث ان كلا منهما يمثل ترتيبا مختلفا للحرفين

التوافيق:

يقصد بالتوافيق بانها عدد طرق الاختيار غير المرتب التي يمكن تكوينها من عدة اشياء باخذها كلها او بعضها ويرمز لها ب C_r^n وحيث

$$C_r^n = n! / r!(n-r)! , n > r$$

ملاحظات/ $C_0^n = 1$, $C_1^n = n$, $C_n^n = 1$, $C_{n-1}^n = n$

مثال/بكم طريقة يمكن اختيار لجنة تحتوي على اربع طلاب من صف فيه عشرون طالب؟

$$C_4^{20} = 20! / 4!(20-4)! \quad \text{الحل/}$$

$$= 20(19)(18)(17)(16)! / 4(3)(2)(1) (16)!$$

$$= 20(19)(18)(17) / 4(3)(2)(1) = 4845$$

قواعد اساسية :

1-اذا كان عدد الطرق الممكنة لوقوع الحادث E_1 هو (n) وان عدد الطرق الممكنة لوقوع الحادث E_2 هو (m) وكان E_1, E_2 حادثان متنافيان فان عدد الطرق الممكنة لوقوع الحادث E_2 او E_1 هو $(n+m)$ من الطرق

مثال / ورق اللعب يتكون من (52) ورقة فورقة () يمكن ان تحدث ب(13) طريقة وان ورقة () يمكن ان تحدث ب (13) طريقة فعند سحب ورقة عشوائية من اوراق اللعب فان عدد الطرق الممكنة لاختيار () او () $= 13+13 = 26$

2-اذا كان عدد الطرق الممكنة لوقوع الحادث E_1 هو (n) وان عدد الطرق الممكنة لوقوع الحادث E_2 هو (m) وكان E_1, E_2 حادثان مستقلان فان عدد الطرق الممكنة لوقوع الحادث E_2 او E_1 هو $(m \times n)$ من الطرق

مثال/اذا سحبت ورقتان من مجموعة اوراق اللعب بحيث ان احدهما () والاخرى () فان هناك $13 \times 13 = 169$ طريقة لعمل ذلك

مثال/ صندوق به (6) كرات حمراء و(4) كرات بيضاء و(2) سوداء ما عدد طرق اختيار (5) كرات بحيث تكون (2) منها بيضاء و(3) حمراء.

الحل/عدد طرق اختيار (3)كرات حمراء هي

$$C3^6 = \frac{6!}{3!(6-3)!} = \frac{6(5)(4)(3)(2)(1)}{3!(3)!} \\ = \frac{6(5)(4)}{3(2)1} = 20$$

وعدد طرق اختيار (2)كرات بيضاء هي

$$C2^4 = \frac{4!}{2!(4-2)!} = \frac{4(3)(2)(1)}{2!(2)!} = 6$$

اذن عدد الطرق لاختيار (3)كرات حمراء (2) بيضاء هو

$$C3^6 * C2^4 = 20 * 6 = 120$$

تعريف الاحتمالية : Probability

هي نسبة الجزء للكل او هي دالة المنطلق لها الحادث (مجموعة الحوادث
المستقر لها

وهي دالة منطلقها مجموعة جزئية من فضاء العينة (الحوادث) والمستقر لها
مجموعة جزئية من الاعداد الحقيقية

فاذا كان عدد عناصر مجموعة فضاء العينة (m) و (n) مجموعة عناصر الحادث (E)
(وتسمى الحالات المواتية Favourable

$$P(E) = \frac{n}{m} = \text{عدد الحالات الممكنة} / \text{عدد الحالات المواتية للحادث}$$

$$, n < m$$

احتمالية الحادث المكمل (احتمالية عدم ظهور هذا الحادث) أي فشله

$$P(E^c) = 1 - P(E) = 1 - \frac{n}{m} = \frac{m-n}{m}$$

مثال/ جد احتمالية وجود صورتان اذا رمي قطعة نقود ثلاث مرات

$$\text{الحل/ } S = [HHH, HHT, HTH, THH, HTT, THT, TTH, TTT]$$

$$A = [HHT, HTH, THH]$$

$$P = \frac{n}{m} = \frac{3}{8}$$

مثال/ عند رمي زهر كم احتمال فردي ؟

$$S = [1, 3, 5],$$

مثال/ صندوق يحتوي على (3) كرات بيضاء و(5)كرات سوداء و(2)حمراء فاذا سحبت كرة عشوائية من هذا الصندوق ما هو احتمال ان تكون :

أ-سوداء ب- غير سوداء ج -غير حمراء

الحل/ احتمال الحصول على كرة سوداء هو $P(E)$

$$P(E) = n/m = 5/10 = 1/2$$

احتمال الحصول على كرة غير سوداء هو $P(E^c)$

$$P(E^c) = 1 - 1/2 = 1/2$$

احتمال الحصول على كرة غير حمراء هو $P(R^c)$

$$P(R^c) = 1 - P(R) = 1 - 2/10 = 1 - 1/5 = 4/5$$

بعض خواص الاحتمالية :

1- اذا رمزنا لاحتمال حدوث حادث بالرمز $P(E)$ واحتمال عدم حدوث هذا الحادث

$$P(E) + P(E^c) = 1 \quad \text{فان } (P(E^c))$$

مثال/ صندوق يحتوي على (6) كرات بيضاء و(4)كرات سوداء و(5)حمراء فاذا سحبت منه كرة واحدة عشوائيا ما هو درجة احتمال ان تكون هذه الكرة :

أ-بيضاء ب- غير بيضاء

$$P(W) = 6/15 \quad \text{الحل/}$$

$$P(W^c) = 1 - 6/15 = 9/15$$

$$P(W) + P(W^c) = 6/15 + 9/15 = 15/15 = 1 \quad \text{ومن هذا يتضح ان}$$

خاصية ثانية :

ان درجة احتمال أي حادث تتراوح بين الصفر والواحد أي $0 < P(E) < 1$

خاصية ثالثة:

اذا كانت E_1, E_2, \dots, E_n هي عناصر او نقاط فضاء العينة فان مجموع درجات احتمالاتها = 1

$$P(E_1) + P(E_2) + \dots + P(E_n) = 1$$

$$P(E_i)=1$$

مثال/ صندوق يحتوي على (5) كرات بيضاء و(9)كرات سوداء و(3)خضراء فاذا سحبت

كرة عشوائيا فان احتمال ان تكون بيضاء هو: $P(E1)= 5 / 20$

احتمال ان تكون سوداء هو: $P(E2)= 9 / 20$

احتمال ان تكون خضراء هو: $P(E3)=6 / 20$

$$P(E1)+P(E2)+P(E3)= 5/20 +9/20 + 6/20 =20 /20 = 1$$

تمارين

1-طالب يجيب على(8)اسئلة من مجموع (10) اسئلة كم طريقة مختلفة للاجابة على الاسئلة؟

2-ما عدد طرق الاختيار التي يمكن الحصول عليها لاختيار لجنة مؤلفة من (5)اشخاص من مجموع (9) اشخاص؟

3-بكم طريقة يمكن ان يجلس ثلاثة اولاد وبنتان في صف؟

بفرض عدم السماح بالتكرار كم عدد مكون من ثلاثة ارقام يمكن تكوينه من الارقام 2,3,5, 6,7,9, ؟

4-سحبت قطرة دم من شخص يرغب في فحص صنف دمه ما هو احتمال ان يكون صنف دم هذا الشخص AB؟

5- صندوق يحتوي على (7) كرات حمراء و(5)كرات بيضاء و(8)خضراء فاذا سحبت منه كرة واحدة عشوائيا فان احتمال ان تكون هذه 1-بيضاء 2-غير بيضاء؟

قوانين الاحتمالية : Laws of probability

نفرض ان هناك حادثان $E1, E2$ فالتعابير التالية يقصد بها مايلي:

احتمال وقوع الحدث $E1$ او الحدث $E2$ $P(E1+E2) =$

أي احتمال وقوع ايا منهما فقط

احتمال وقوع الحادث $E1$ والحادث $E2$ معا $P(E1E2) =$

احتمال حدوث $E2$ علما بان الحادث $E1$ قد وقع $P(E2/E1)$ ويسمى الاحتمال الشرطي (Conditional Probability)؟

والحوادث ما يلي :أ- قانون الجمع (Addition law)

1- اذا كانت الاحداث متنافية ، أي اذا كانت $E1$ و $E2$ حادثان متنافيان فان احتمال حدوث أي منهما ($E1$ و $E2$) هو حاصل جمع احتمال كل منهما أي

$$P(E1+E2) = P(E1) + P(E2)$$

مثال/ صندوق يحتوي على (6) كرات بيضاء و (3) كرات سوداء و (5) زرقاء فاذا سحبت منه كرة واحدة فان احتمال ان تكون هذه بيضاء او سوداء؟

$$P(E1+E2) = P(E1) + P(E2) \quad \text{الحل/}$$

$$6/14 + 3/14 = 9/14$$

مثال/ في حالة رمي زهر نرد (زار) فما هو احتمال الحصول على عدد فردي؟

$$P(1+3+5) = P(1) + P(3) + P(5) = 1/6 + 1/6 + 1/6 = 3/6 = 1/2$$

2- اذا كانت الاحداث غير متنافية ، أي اذا كانت $E1$ و $E2$ حادثان غير متنافيان فان احتمال حدوث أي منهما ($E1$ و $E2$) هو حاصل جمع احتمال كل منهما مطروحا منه احتمال حدوثهما معا أي

$$P(E1+E2) = P(E1) + P(E2) - p(E1E2)$$

مثال/ في احدى الكليات 25% من الطلبة راسب بالرياضيات و 15% ومن الطلبة راسب في الكيمياء و 10% راسب في كلا الرياضيات والكيمياء فاذا انتخب طالب منهما عشوائيا فما هو احتمال ان يكون راسبا في الرياضيات او الكيمياء؟

الحل/ نرسم للرياضيات ب M وللكيمياء ب C

$$p(M+C) = p(M) + p(C) - p(MC) = 0.25 + 0.15 - 0.10 = 0.30$$

مثال/ اذا رمي زار مرة واحدة فما هو احتمال ظهور عدد يكون فردي او يقبل القسمة على 3؟

الحل/ نفرض $E1$ الحادث الفردي وعدد حالاته الممكنة (3) وهي (1,3,5)

نفرض الحادث $E2$ يقبل القسمة على (3) وعدد حالاته الممكنة (2) وهي (3,6)

$$P(E1) = 3/6 , \quad P(E2) = 2/6$$

يمثل $E1E2$ الحادث فردي ويقبل القسمة على 3 هو فقط 3 وعدد الحالات

$$P(E1E2) = 1/6 \quad \text{الممكنة} = 1$$

$$P(E1+E2) = P(E1) + P(E2) - P(E1E2)$$

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

ب- قانون الضرب :

3- اذا كانت الاحداث مستقلة ، أي اذا كانت $E1$ و $E2$ حادثين مستقلين فان احتمال حدوثهما معا هو حاصل ضرب كل منهما أي

$$P(E1E2) = P(E1) P(E2)$$

مثال/ عند رمي قطعتي نقود ما هو احتمال الحصول على صورة في كليهما؟

$$P(H1) \text{ الحل/ احتمال الحصول على صورة من القطعة الاولى}$$

احتمال الحصول على اصورة من القطعة الثانية $P(H2)$

$$P(H1H2) = P(H1) P(H2) = (1/2)(1/2) = 1/4 \text{ اذن}$$

مثال/ اذا كان احتمال اصابة الطائرة الاولى (A) لهدف معين يساوي $(1/4)$ واحتمال اصابة الطائرة الثانية (B) لنفس الهدف يساوي $(1/6)$ فاذا كان عمل كل منهما مستقلا عن الاخر واطلقت كل من الطائرتين قنبلة في آن واحد تجاه الهدف فما هو احتمال أ - انتصيب الطائرتان معا الهدف ب- ان لا تصيبا الهدف؟

$$P(AB) = P(A) P(B) = (1/4)(1/6) = 1/24 \text{ الحل /}$$

ب- احتمال الاولى ان لا تصيب الهدف

$$P(A^c) = 1 - P(A) = 1 - 1/4 = 3/4 \text{ -}$$

احتمال الثانية ان لا تصيب الهدف

$$P(B^c) = 1 - P(B) = 1 - 1/6 = 5/6 \text{ -}$$

احتمال ان لا تصيبا الهدف

$$P(A^c * B^c) = P(A^c) P(B^c) = (3/4)(5/6) = 15/24 \text{ -}$$

4- اذا كانت الاحداث غير مستقلة ، أي اذا كانت $E1$ و $E2$ حادثان غير مستقلين فان احتمال حدوثهما معا يساوي حاصل ضرب احتمال وقوع الحادث الاول في احتمال وقوع الحادث الثاني بشرط حدوث الاول أي

$$P(E1E2) = P(E1) P(E2 / E1)$$

مثال / صندوق به (5) كرات حمراء و(3) سوداء فاذا سحبت كرتان سوية (او سحبت كرتان على التوالي بدون ارجاع الكرة الاولى الى الصندوق) ماهو احتمال ان تكون كلتاها سوداء ؟

الحل/ احتمال الحصول على كرة سوداء في السحبة الاولى هو:

$$P(E1) = 3/8$$

اما في السحبة الثانية (بدون ارجاع الكرة المسحوبة الى الصندوق)

فان احتمال ان تكون الكرة سوداء هو :

$$P(E2/E1) = 2/7$$

$$P(E1E2) = P(E1) P(E2/E1) = (3/8)(2/7) = 6/56$$

الاحتمال الشرطي: Conditional probability:

اذا كان الحادث A قد وقع فعلا والمطلوب احتمال وقوع الحادث B بحيث ان B يعتمد على A فان الاحتمال الشرطي علما ان (A,B غير مستقلين) هو كالاتي:

$$P(B/A) = P(AB) / P(A) , P(A) > 0$$

مثال/ رمينا قطعة نقود مرتان وارادنا ان نحسب احتمال حصولنا على صورة في الرمية الاولى مع العلم اننا لم نحصل على كتابة مرتين ؟

الحل/ الحادث بالرمية الاولى صورة $S = [HH, TH, HT, TT]$

$$P(A) = 2/4$$

الحادث عدم الحصول على كتابتين $B = [HH, TH, HT]$

$$P(B) = 3/4$$

$$P(A/B) = P(AB) / P(B) = (2/4) / (3/4) = 2/3$$

مثال / تم تصنيف احدى الافراد وفق الجدول التالي:

الجنس	له وظيفة E	ليس له وظيفة V	المجموع
M ذكور	460	40	500
F اناث	140	260	400
المجموع	600	300	900

ثم اختيار فرد بصورة عشوائية فما احتمال ان يكون ذكر وله وظيفة ؟

$$\begin{aligned} \text{الحل/} \quad P(M/E) &= P(ME) / P(E) = (460/900) / (600/900) \\ &= 460 / 600 = 0.76 \end{aligned}$$

مثال واجب/ نفس المثال السابق ما احتمال ان يكون انثى وليس لها وظيفة؟

الفصل الرابع :

" Probability Distribution " التوزيع الاحتمالي

المتغير العشوائي : Random Variable

هو دالة ذات قيمة حقبقة معرفة على مجموعة فضاء العينة حيث S تمثل المنطلق ومستقر الدالة مجموعة جزئية من الاعداد الحقيقية ويرمز له باحدى تاتحرف الكبيرة بحيث ان لكل X عدد حقيقي x قيمة المتغير العشوائي.

$$X(w) < x, \quad w \text{ ينتمي لـ } S$$

مثال /رمي قطعة نقود معينة مرتين وكان المتغير x يمثل عدد الصور ؟

$$\text{الحل/} \quad S = [HH, HT, TH, TT]$$

$$X(TT) = 0, \quad X(HT) = X(TH) = 1, \quad X(HH) = 2$$

$$X = 0, 1, 2$$

ويمكن التميز بين نوعين من المتغيرات العشوائية اعتمادا على نوع القيم التي يأخذها المتغير العشوائي وهي :

1- المتغير العشوائي الذي تختلف قيمة الواحدة عن الاخرى بكميات محددة يسمى المتغير العشوائي المتقطع

2- المتغير العشوائي الذي يأخذ أي قيمة ضمن حدود معينة يسمى المتغير العشوائي المستمر $V \in \mathbb{R}$. Continuous .

التوزيع الاحتمالي المتقطع :

هو قانون او جدول يعطي جميع قيم المتغير العشوائي المتقطع مع جميع الاحتمالات المترتبة مع كل قيمة من قيم المتغير المتقطع ويرمز له ب بحيث يحقق الشرطين وهما :

$$1-P(X) > 0$$

$$2- P(X_i) = 1$$

وتسمى $P(X)$ بدالة الاحتمال

مثال / بالنسبة للمثال السابق

احتمال المتغير العشوائي $X=0$ يعني احتمال عدم الحصول على الصورة ؟

$$P(X=0) = 1/4 \quad \text{الحل/}$$

$$P(X=1) = 2/4 = 1/2$$

$$P(X=2) = 1/4$$

$$P(X) = [1/4 , X= 0, 2 \quad \text{اذن}$$

$$1/2 , X= 1$$

$$0 , o.w$$

او

X	0	1	2
P(X)	1/4	1/2	1/4

مثال / دالة الكتلة الاحتمالية لرمي الزار للمتغير العشوائي X الذي يمثل عدد النقاط عند رمي الزار . ؟

$$P(X) = 1/6 , X= 1, 2, 3, 4, 5, 6 \quad \text{الحل/}$$

$$= 0 , o.w$$

X	1	2	3	4	5	6
P(X)	1/6	1/6	1/6	1/6	1/6	1/6

التوزيع الاحتمالي المستمر :

هو التوزيع الذي يأخذ فيه المتغير العشوائي قيما بين حدين ودالته $f(x)$ تكون موجبة لجميع قيم x بين $x < \infty$ - تسمى $f(x)$ بدالة الكثافة الاحتمالية ومن خواص هذه الدالة هو :

$$1-f(x) > 0 \quad 2- \int_{-\infty}^{\infty} f(x) = 1$$

مثال / من تجربة رمي زارين جد التوزيع الاحتمالي لمجموع النقاط التي تظهر على الوجهين فان $X= 2,3,4,\dots,12$

ان عدد عناصر فضاء العينة هو

$$S= [(1,1), (1,2), \dots, (1,6)$$

$$(2,1), (2,2), \dots, (2,6)$$

$$(6,1), (6,2), \dots, (6,6)$$

عندما فالعنصر هو $(1,1)$ فقط أي ان : $P(X=2) = 1/36$

وعندما فالعناصر هي $(1,2)$ و $(2,1)$ فقط أي ان : $P(X=3) = 2/36$

وهكذا يمكن ايجاد الاحتمال عند كل قيمة من قيم x وبذلك يكون التوزيع الاحتمالي للمتغير العشوائي x هو كما يلي :

X	2	3	4	5	6	7	8	9	10
P(X)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36

ومجموع هذه الاحتمالات = 1

التوقع الرياضي :

إذا كان X متغير عشوائي بدالة كثافة احتمالية أو كتلة احتمالية $f(X), P(X)$ وكانت $g(X)$ دالة أخرى بدلالة المتغير العشوائي فالتوقع الرياضي لهذه الدالة يعرف بالصيغة التالية ويرمز له بـ $E(g(X))$

$$E[g(x)] = \sum g(x) P(x) \quad \text{في حالة المتغير العشوائي المتقطع}$$

$$E[g(x)] = \int g(x) f(x) dx \quad \text{في حالة المتغير العشوائي المستمر}$$

بعض الحالات الخاصة للتوقع الرياضي:

أ_ المعدل (Mean) : إذا كانت $g(x) = X$ فالتوقع الرياضي في هذه الحالة يسمى القيمة المتوقعة للمتغير العشوائي ويرمز له بـ M

$$E(X) = M = \sum x P(X) \quad \text{متغير عشوائي متقطع}$$

$$E(x) = M = \int x f(x) dx \quad \text{متغير عشوائي مستمر}$$

ويعني القيمة المتوقعة هي في الحقيقة قيمة المتوسط النظري للمجتمع M أي $E(X)=M$

ب_ التباين (Variance)

إذا كانت الدالة $g(X) = (X-M)^2$ فالتوقع الرياضي في هذه الحالة تباين المتغير العشوائي X ويرمز له بـ $S^2 = E(X-M)^2$

$$S^2 = E(X^2) - (E(X))^2 \quad \text{أو}$$

خواص التوقع الرياضي:

إذا كان X متغير عشوائي بدالة كثافة (كتلة) و $C1, C2$ ثوابت فان :

$$1-E(C1) = C1$$

$$2-E(C1X) = C1E(X)$$

$$3-E(C1X+C2) = C1E(X) + C2$$

مثال / إذا كان متغير عشوائي بدالة كثافة احتمالية : $0 < X < 1$, $F(X)=2X$ احسب معدل وتباين المتغير العشوائي X ؟

$$E(X) = M = \int Xf(X) dx = \int 2X^2 dx = 2/3 \quad /الحل$$

$$S^2 = \text{Var}(X) = \int (X-M)^2 f(X) dx = \int (X-2/3)^2 (2x) dx$$

$$(2X^3 - 8/3 X^2 + 8/9 X) dx = 1/2 x^4 - 8/9 X^3 + 4/9 X^2$$

$$1/2 - 8/9 + 4/9 = 1/2 - 4/9 = 1/18$$

مثال / من التوزيع الاحتمالي المتقطع الاتي احسب متوسط المجتمع وتباينه ؟

X	0	1	2	3	4
P(X)	0.15	0.15	0.35	0.25	0.10

/الحل

$$M = E(X) = \sum X_i P(X_i) =$$

$$(0)(0.15) + (1)(0.15) + (2)(0.35) + (3)(0.25) + (4)(0.10) = 2$$

$$S^2(X) = E(X - M)^2 = (0-2)^2(0.15) + (1-2)^2(0.15) + (2-2)^2(0.35) + (3-2)^2(0.25) + (4-2)^2(0.10) = 1.4$$

مثال/ اذا علمت ان $P(X) = C^3(X) (1/2)^3$, $X = 0, 1, 2, 3$ جد

$$E(X^2 + 1) ?$$

الحل / نجد التوزيع الاحتمالي للمتغير العشوائي كالاتي :

X	0	1	2	3
P(X)	1/8	3/8	3/8	1/8

$$E(X^2 + 1) = E(X^2) + 1 = \sum X_i^2 P(X_i) + 1$$

$$= (0)^2(1/8) + (1)^2(3/8) + (2)^2(3/8) + (3)^2(1/8) = 4$$

مثال/ اذا كان $F(X) = 1/2$, $1 \leq X \leq 3$

أ_ برهن ان $f(X)$ هي دالة كثافة احتمالية.

$$ب_ جد (2 < X < 2.5) P$$

$$ج_ جد (x \leq 1.6) P$$

$$\int f(X)dx = \int 1/2 dx = 1/2x \Big| = 1/2(3-1) = 1 \quad \text{أ- الحل /}$$

$$P(2 < X < 2.5) = \int F(X) dx = 1/2 \int dx = 1/2 x \Big| = 1/4 \quad \text{ب-}$$

$$P(X < 1.6) = \int f(X) dx = 1/2 \int dx = 1/2 X \Big| = 0.3 \quad \text{ج-}$$

خواص التباين: يرمز للتباين ب ∂X^2 او $\text{Var}(X)$ فاذا كان a, b ثوابت

$$1- \text{V}(a) = 0 \quad \text{فان:}$$

$$2- \text{V}(aX) = a^2 \text{V}(x) = a^2 \partial x^2$$

$$3- \text{V}(X \pm a) = \text{V}(X) = \partial x^2$$

$$4- \text{V}(ax \pm b) = a^2 \text{V}(x) = a^2 \partial x^2$$

مجموعة تمارين /

1- احسب توقع وتباين المتغير العشوائي الذي دالته الاحتمالية كما يلي :

$$a- P(X) = X / 15 , \quad X = 1, 2, 3, 4, 5$$

$$b- f(X) = 3x^2 , \quad 0 < X < 1$$

2- من التوزيع الاحتمالي الاتي :

X	8	12	16	20	24
P(X)	1/8	1/6	3/8	1/4	1/12

جد : a $E(X)$ b $E(X^2)$ c التباين

3- اذا كان : $f(X) = 2X , 0 < X < 1$

أ- برهن ان $f(X)$ هي دالة كثافة احتمالية

ب- جد : $P(1/2 < X < 3/4)$

من التوزيع الاحتمالي الاتي :

X	-2	-1	0	1	2
P(X)	1/8	1/4	1/4	1/4	1/8

$$\text{أ_جد : } \partial x^2 \quad \text{ب_ } E(X^2 - 2X + 3)$$

$$\text{ج_ } V(3X - 4) \quad \text{د_ } V(-8X)$$

التوزيعات الاحتمالية المتقطعة :

توزيع ذي الحدين : Binomial Distribution

لنفرض ان تجربة تكررت (n) من المرات (المحاولات) وان (p) يمثل احتمال النجاح في كل محاولة و (q) يمثل احتمال الفشل في كل محاولة و (q = 1 - p)

اذا فرضنا X يمثل عدد المرات التي يحدث فيها النجاح والقيم الممكنة هي (0,1,2,...n) فان X متغير عشوائي له توزيع ذي الحدين ودالة كتلة احتمال هي

$$P(X) = Cx^n P^x (1-P)^{n-x}, X = 0,1,2,\dots,n$$

$$Cx^2 = n! / x! (n-x) ! \quad \text{حيث:}$$

وتوزيع ذي الحدين يكون بالمعلمة P , n أي ان

$$X \approx B(n, p) . M = np , \partial x^2 = V(x) = npq$$

مثال / اذا كانت نسبة الوحدات الرديئة المنتجة في معمل ما هي ()

وقد اخذت عينة من خمسة وحدات جد احتمال:

1-كون عدد الوحدات التالفة او الرديئة تقل عن الربع.

ان نجد على الاقل ثلاثة وحدات سالحة في العينة

/الحل

$$P = 10 / 100 = 1 / 10 , q = 1 - p = 1 - 1/10 = 9/10 , n = 5$$

$$B(5, 1/10) \sim X$$

$$P(X) = C x^5 (1/10)^x (9/10)^{5-x}, X=0,1,2,3,4,5$$

$$1- P(X < 5/4) = P(X < 1.25) = P(0) + P(1)$$

$$C 0^5 (1/10)^0 (9/10)^5 + C 1^5 (1/10) (9/10)^4 =$$

$$(9/10)^5 + 1/2 (9/10)^4 = (9/10)^4 [9/10 + 1/2]$$

$$= (9/10)^4 [14/10]$$

$$2- P(X \geq 3) = 1 - P(X < 3) = 1 - [P(0) + P(1) + P(2)]$$

$$1 - [(9/10)^4 + C 2^5 (1/10)^2 (9/10)^3 =$$

مثال/ عائلة مكونة من اربعة اطفال اوجد دالة احتمال بحيث يكون لديهم ولدين؟

الحل/ نفرض عدد الاولاد X

$$P = 1/2, q = 1/2, X = 0, 1, 2, 3, 4$$

$$B(4, 1/2) \sim X$$

$$P(X) = C x^4 (1/2)^x (1/2)^{4-x}$$

$$P(X=2) = C 2^4 (1/2)^2 (1/2)^2 = 3/8$$

مثال/ اذا كان لدينا X يتوزع حسب توزيع ذي الحدين أي :

$$X \sim B(n, p) \text{ و لدينا } \partial x^2 = 4/3, M=2 \text{ جد } n, p ?$$

$$M = n p = 2 \text{1 / الحل}$$

$$\partial x^2 = n p q = 4/3 \text{2}$$

نعوض قيمة n في المعادلة 2 ←

$$\partial x^2 = n p q = n p (1-p) = 4/3$$

$$(2/p) p (1-p) = 4/3 = 2(1-p) = 4/3$$

$$P=1/3 \rightarrow n=6$$

$$X \sim B(6, 1/3) \quad \text{اذن}$$

توزيع الحوادث النادرة:

توزيع بواسون

يعرف توزيع بواسون بتوزيع الحوادث النادرة حيث ان احتمالية وقوع حادث صغير جدا وعدد المحاولات كبيرة فالمتغير العشوائي الذي يمثل حصول على محاولة ناجحة يكون له توزيع بواسون

$$P(X) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x=0,1,2,\dots, \lambda > 0$$

حيث ان λ هو متوسط (معدل) النجاحات

$$X \sim P(\lambda)$$

اما الحالات التي يستخدم فيها هذا التوزيع :

- 1- الحوادث التي تحدث عن طريق الصدفة في زمن معين
- 2- الاخطاء المطبعية في صفحة معينة من كتاب
- 3- عدد البكتريا في محلول معين الخ

مثال/ اذا علمت ان معدل السيارات المتعطلة في شارع معين هو () سيارات باليوم الواحد ماهي احتمالية ان تتعطل سيارة واحدة في الشارع المذكور ما بين الساعة الثامنة والنصف والتاسعة صباحا؟

الحل/ معدل السيارات المتعطلة خلال اليوم الواحد هي (3) سيارات هو = $3/24$

معدل السيارات المتعطلة خلال $1/2$ ساعة هو

$$3/24 (1/2) = 3/48$$

$$X \sim P(3/48)$$

$$P(X) = \frac{(3/48)^x (e^{-(3/48)})}{x!}, \quad x = 0, 1, 2, \dots$$

$$P(X=1) = (3/48) (e^{-3/48})$$

مثال / اذا كان عدد الاخطاء المطبعية في كتاب يحتوي على (600) صفحة هو (400) خطأ جد احتمال ان صفحة واحدة تحتوي على ثلاثة اخطاء مطبعية؟

الحل / وجود خطأ واحد في الصفحة (1 / 600)

وجود خطأ في الكتاب كله (400)

$$\lambda = (1/600)(400) = 2/3 \quad \text{معدل بصورة عامة:}$$

$$X \sim P(2/3)$$

$$P(X) = \frac{(2/3)^x (e^{-(2/3)})}{x!}, \quad x = 0, 1, 2, \dots$$

$$P(X=3) = \frac{e^{-(2/3)}(2/3)^3}{3!} = \frac{e^{-(2/3)}(2/3)^3}{6}$$

مثال /

لوحظ في منطقة معينة ان من بين كل () شخص هنالك شخص واحد مصاب بمرض معين ما هو احتمال ان من بين () شخص هنالك

أ- ثلاثة اشخاص مصابين بهذا المرض

ب- اكثر من شخصين مصابين بهذا المرض

$$\lambda = np = (2000) (0.001) = 2 \quad \text{الحل /}$$

$$X \sim P(2)$$

$$P(X) = \frac{(2^x)(e^{-2})}{x!}, \quad x = 0, 1, \dots$$

$$P(X=3) = \frac{(2^3)(e^{-2})}{3!} = 0.18045$$

$$P(X > 2) = 1 - P(X \leq 2) = 1 - [p(x=0) + p(x=1) + p(x=2)]$$

$$1 - [(2^0)e^{-2}/0! + (2^1)e^{-2}/1! + (2^2)e^{-2}/2!]$$

$$= 1 - (2^0e^{-2}/0! + 2^1e^{-2}/1! + 2^2e^{-2}/2!)$$

$$=1 - (e^{-2} + 2e^{-2} + 2e^{-2})$$

$$=1 - 0.67668 = 0.32332$$

تمارين

احد الفرق الرياضية يكون احتمال الفوز لها (2/3) فاذا تم اللعب (4) مرات ما هو احتمال الفوز

-مرتين بالضبط

- على الاقل مرة واحدة

- اكثر من نصف مرات اللعب

اذا كانت نسبة الوفاة نتيجة الاصابة بمرض معين هي (25%) اختير سبعة اشخاص مصابين بهذا المرض جد احتمال ،

-على الاقل ثلاث منهم سيموتون

-اثنان على الاكثر سيشفون

اذا كان (4%) من انتاج معمل هو سلع رديئة جد احتمال وجود سلعتين على الاقل رديئة في عينة تتكون من (100) سلعة .

شركة تأجير سيارات تملك اربع سيارات فاذا افترضنا معدل الطلبات على هذه السيارات هو (2.5) باليوم جد احتمال ،

-ان تتلقى الشركة عدد من الطلبات اكثر من عدد سيارات الشركة .

- ان يمر يوم دون ان تتلقى هذه الشركة أي طلب على هذه السيارات.

التوزيعات المستمرة: Continuous Distributions

التوزيع الطبيعي: Normal Distribution

اذا كان المتغير العشوائي X يتوزع توزيعا طبيعيا وله وسط حسابي μ وتباين θ وان معادلة المنحني الطبيعي هي :

$$f(x) = \frac{1}{\sqrt{2\pi\theta}} e^{-(x-\mu)^2/2\theta^2} , \quad -\infty < x < \infty , \quad \theta > 0$$

خصائص التوزيع الطبيعي :

- التمثيل البياني له منحنى يشبه الجرس و متمائل بالنسبة للوسط
- يتساوى الوسط و الوسيط و المنوال و يقع في المركز
- المنحنى متصل
- يقترب المنحنى من المحور x ولكنه لا يمسه
- المساحة تحت المنحنى تساوي 1 او 100% .

يتصف التوزيع الطبيعي الذي وسطه μ وانحرافه المعياري σ بالخصائص الآتية

- يقع 68% تقريبا من البيانات ضمن الفترة $\mu - \sigma$ ، $\mu + \sigma$
- يقع 95% تقريبا من البيانات ضمن الفترة $\mu - 2\sigma$ ، $\mu + 2\sigma$
- يقع 99% تقريبا من البيانات ضمن الفترة $\mu - 3\sigma$ ، $\mu + 3\sigma$

$$X \sim N(\mu, \sigma^2)$$

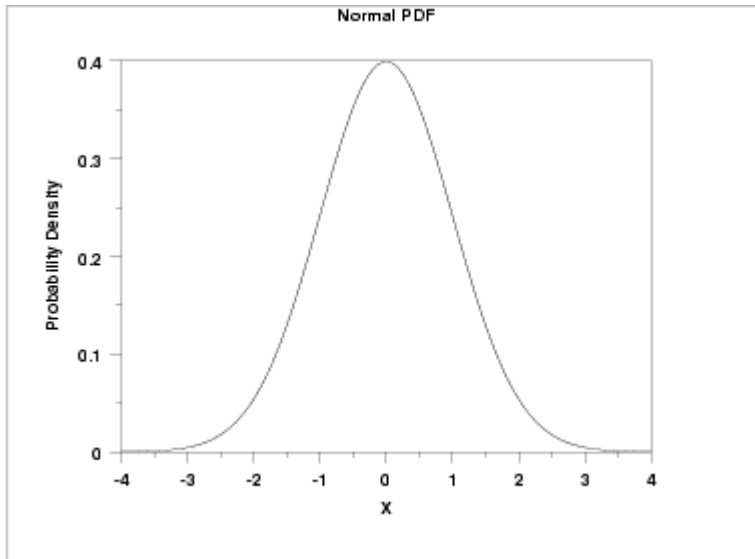
التوزيع الطبيعي : Normal Distribution

يقال للمتغير العشوائي X بأنه له التوزيع الطبيعي القياسي

Standard Normal Distribtuion بوسط حسابي يساوي صفر وتباين يساوي واحد اذا

كانت دالة الكثافة الاحتمالية للمتغير العشوائي Z بالصيغة التالية

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty$$



مجالات استخدام التوزيع الطبيعي :

-الاحصاءات الوبائية كالوفيات والولادات في فترة معينة وفي كل مكان معين

- الظواهر الاجتماعية كالزواج والطلاق تخضع لتوزيع قريب من التوزيع الطبيعي وكذلك نسبة الدخول ومستوى الانتاج

- المقاييس النفسية والعقلية كالذكاء والقابليات العقلية

بما ان $f(x)$ دالة كثافة احتمالية $P.D.F$ لذا فان $\int_{-\infty}^{\infty} f(x)dx = 1$

$$\underline{X \sim N(0, 1)}$$

والان اصبح من السهل تحويل المتغيرات العشوائية X التي تتوزع توزيعا طبيعيا الى المتغيرات عشوائية Z والتي تتوزع توزيع طبيعي قياسي وذلك بالطريقة التالية :

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

مثلا: اذا كانت X بين الحدين x_1, x_2 فان Z لها ستكون بين Z_1, Z_2 حيث ان

$$Z_1 = \frac{x_1 - \mu}{\sigma} \quad , \quad Z_2 = \frac{x_2 - \mu}{\sigma}$$

ملاحظة : اذا كان متغير عشوائي له التوزيع الطبيعي فان

$$P(X < x) = N(x)$$

باستخدام جدول التوزيع الطبيعي القياسي

امثلة :

- $\underline{X \sim N(3, 4)}$ فان

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot 2} e^{-(x-3)^2/2(4)}$$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-(x-3)^2/8} \quad \text{اذن}$$

- $\underline{X \sim N(-2, 9)}$ فان

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-(x+2)^2/18}$$

$$\text{اذن } f(x) = \frac{1}{\sqrt{\pi}} e^{-(x-3)^2} \quad \text{جد } \mu, \sigma$$

الحل :

$$(x - \mu)^2 = (x-3)^2 \quad \text{بما ان}$$

$$\mu = 3 \quad \text{اذن}$$

$$f(x) = \frac{1}{\sqrt{2\pi} \sigma} = f(x) = \frac{1}{\sqrt{\pi}} \rightarrow \frac{1}{\sqrt{2} \sigma} = 1$$

$$\rightarrow \sigma = \frac{1}{\sqrt{2}}$$

$$\rightarrow \sigma^2 = \frac{1}{2}$$

مثال :

إذا كان الوسط الحسابي لتوزيع طبيعي هو $(\mu = 25)$ والانحراف القياسي (المعياري) له هو $(\sigma = 10)$ جد قيمة z_1, z_2 بحيث ان :

$$P(45 < X < 62) = P(z_1 < Z < z_2)$$

$$P(z_1 < 0.5) = 0.933, P(z_2 < 12) = 0.8849 \quad \text{علما بان :}$$

الحل:

$$Z_1 = \frac{x_1 - \mu}{\sigma} = \frac{45 - 50}{10} = -0.5$$

$$Z_2 = \frac{x_2 - \mu}{\sigma} = \frac{62 - 50}{10} = 1.2$$

$$P(45 < X < 62) = P(-0.5 < Z < 1.2)$$

$$P(-0.5 < Z < 1.2) = P(Z < 1.2) - P(Z < -0.5)$$

$$= 0.8849 - 0.3085 = 0.5764$$

مثال:

إذا كان X له التوزيع الطبيعي $X \sim N(3, 4)$ جد $P(X \leq 2)$ ؟

الحل:

$$P(X \leq 2) = P\left(\frac{x - \mu}{\sigma} \leq \frac{2 - 3}{2}\right) = P\left(Z \leq -\frac{1}{2}\right)$$

$$= N\left(-\frac{1}{2}\right) = 1 - (0.5) = 1 - 0.691 = 0.309$$

مثال :

إذا كانت درجات الحرارة خلال شهر نيسان في مدينة ما تتوزع توزيعا طبيعيا بوسط حسابي هو 20°C وانحراف معياري 2°C جد احتمال ان تكون درجة الحرارة للاحد الايام :

- اقل من 18 c
- اكثر من 22 c
- بين 18 c و 22 c

الحل :

$$X \sim N(20, 4)$$

$$P(X < 18) = P\left(\frac{x-\mu}{\sigma} < \frac{18-20}{2}\right)$$

$$P\left(Z < \frac{-2}{2}\right) = N(-1) = 1 - N(1) = 1 - 0.841 = 0.159$$

$$P(X > 22) = 1 - P(X \leq 22) = 1 - P\left(Z \leq \frac{22-20}{2}\right)$$

$$= 1 - N(1) = 1 - 0.841 = 0.159$$

$$P(18 < X < 22) = P(X < 22) - P(X < 18)$$

$$P\left(Z < \frac{22-20}{2}\right) - P\left(Z < \frac{18-20}{2}\right)$$

$$= N(1) - N(-1) =$$

$$N(1) - 1 + N(1) = 2N(1) - 1$$

$$2(0.841) - 1 = 0.682$$

ملاحظة:

$$F(-X) = 1 - F(X)$$

تمارين :

- 1- اذا كان متوسط انتاج الدونم من الذرة الطفراء هو 800 كغم وبانحراف قياسي قدره 40 كغم وعلى فرض ان كمية المحصول يتتبع التوزيع الطبيعي . ما هو احتمال ان نباتا يعطي محصولا بين 778-834 كغم ؟

$$2- \text{ اوجد } X \sim N(2, 25)$$

$$P(X < 0)$$

$$P(0 < X < 10)$$

$$P(|X| < 10)$$

- 3- اذا كان اوزان الف طفل عند الولادة يتوزع توزيع طبيعي يوسط حسلي (8.5) وانحراف قياسي (معياري) (0.5) باون احسب عدد الاطفال باوزان

$$\text{اقل من } 9.7$$

$$\text{بين } 7.2 \text{ و } 8.5$$

Chapter 5 : الفصل الخامس :

Tests of Hypotheses اختبار الفرضيات

Statistical Hypothesis : الفرضية الاحصائية :

هي عبارة عن ادعاء او تصريح (قد يكون صائبا او خطأ) حول معلمة او اكثر لمجتمع او لمجموعة مجتمعات.

عندما نأخذ عينة من مجتمع ما لمعرفة فرضية معينة نستخدم جميع المعلومات من العينة للوصول الى قرار بقبول او رفض الفرضية الاحصائية ،

فالاحصائي او الباحث يحاول دائما ان يضع الفرضية بشكل يأمل ان يرفضها مثلا:

اذا اراد الباحث ان يقارن بين صنفا جديدا من الحنطة مع الصنف المحلي . فانه يضع فرضية فحواها بانه لا يوجد فرق معنوي او جوهري بين الصنفين وكذلك اذا اراد ان يبرهن بان طريقة جديدة من طرق التدريس احسن من غيرها فانه يضع فرضية تقول بعدم وجود فرق بين طرق التدريس هذه وهكذا .

ان الفرضية التي يضعها الباحث على امل ان يرفضها تدعى بفرضية العدم

(Null Hypothesis) ويرمز لها ب H_0 وفرضنا لفرضية العدم يقودنا الى قبول فرضية بديلة عنها هذه الفرضية تدعى الفرضية البديلة (Alternative Hypothesis) ويرمز لها ب H_1 .

-مستوى المعنوية Level of Significant

او مستوى الاحتمال Probability Level

او حجم الاختبار Size of the Test

ويعرف مستوى المعنوية :

بانه درجة الاحتمال الذي يرفض به فرضية العدم (H_0) عندما تكون صحيحة او (بعبارة اخرى هو احتمال الوقوع في الخطا من النوع الاول ويرمز لها ب α) يرمز لدرجة الاحتمال α

هذه الدرجة () تحدد من قبل الباحث وهي اما () او () على الاكثر . عندما نأخذ مستوى احتمال () يعني بانه اذا تكررت التجربة لعدد كبير من المرات فمن المحتمل ان نرفض فرضية العدم () بالرغم من انها صحيحة مرة كل () مرة أي ان احتمال الوقوع في الخطا في الاستنتاج من النوع الاول هو () او اقل وان الاستنتاج يكون صائبا وسليم بدرجة ثقة () كما ان مستوى احتمال () يعني بانه فمن المحتمل ان نرفض فرضية العدم () وهي صحيحة خمس مرة في كل () مرة فاحتمال الوقوع في الخطا في الاستنتاج من النوع الاول هو () او اقل بينما

الاستنتاج يكون صائبا وسليم بدرجة ثقة () اما اذا قبلنا فرضية العدم دل ذلك على وجود فرق معنوي .

-المختبر الاحصائي :

عبارة عن متغير عشوائي له توزيع معلوم ويصف المختبر الاحصائي العلاقة بين القيم النظرية للمجتمع والقيم المحسوبة للعينة

-منطقة الرفض

او المنطقة الحرجة :

ان منطقة الرفض هي تلك المنطقة التي اذا وقعت قيمة المختبر الاحصائي داخلها تسبب في رفض فرضية العدم () اما المنطقة الاخرى غير منطقة الرفض تسمى منطقة القبول

() : وهي المنطقة التي اذا وقعت قيمة المختبر الاحصائي داخلها تسبب في قبول فرضية العدم () .

اختبارات تتعلق بالمتوسطات

1-اختبارات تتعلق بمتوسط واحد

الاختبار يتعلق بالفرضية الاتية $H_0: \mu = \mu_0$

حيث μ هو الوسط الحسابي للمجتمع

μ_0 هو قيمة معينة معلومة ، علما ان تباين المجتمع σ^2 معلوم وبعبارة اخرى سنقارن هنا متوسط عينة بمتوسط المجتمع لنرى هل ان العينة تنتمي لهذا المجتمع ام لا؟

فاذا كانت نتيجة الاختبار بالايجاب ففي هذه الحالة يكون متوسط العينة المحسوب لا يختلف اختلافا جوهريا عن متوسط المجتمع .

صياغة فرضيتنا العدم والبديلة :

عند اختيار متوسط المجتمع μ يساوي قيمة معينة μ_0 ضد فرضية القائلة بان μ لا يساوي μ_0 فاننا نكتب الفرضيتان بالشكل الاتي :

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

عند رفض فرضية العدم فان الفرضية البديلة تاخذ مدى من القيم يشمل جميع القيم الممكنة لمعلمة المجتمع

مثلا:

عند اختبار درجة تعبئة معمل تعليب معجون الطماطة التي تزن (250) غم فاننا نضع فرضية العدم (H_0) بان متوسط وزن علبة المعجون هو (250) غم ، اما الفرضية البديلة فهي في هذه الحالة ان متوسط علبة المعجون لا تساوي (250) غم أي :

$$H_0: \mu = 250$$

$$H_1: \mu \neq 250$$

فاذا رفضنا فرضية العدم وقبلنا الفرضية البديلة يعني ذلك ان متوسط وزن العلبة لايساوي (250) غم وهذا معناه ان متوسط الوزن قد يكون اكثر او اقل من (250) غم

مثال آخر :

نفرض ان وزارة الصحة تريد اختبار فاعلية دواء مستحدث عن الدواء الحالي في علاج مرض معين فاذا كان متوسط شفاء المرضى المصابين بهذا المرض بعد تناولهم الدواء الحالي هو (70%) أي ان ($p=0.70$) بينما اذا اعطيت للمرضى الدواء الجديد الذي قيل بانه اكثر فعالية من الدواء الحالي فان (p) في هذه الحالة ستكون اكثر من (70%) أي ($p > 0.70$) أي ان فرضية العدم والبديلة ستكون

$$H_0: p = 0.70$$

$$H_1: p > 0.70$$

واذا كان العكس فان الفرضيتان تكون :

$$H_0: p = 0.70$$

$$H_1: p < 0.70$$

اتخاذ القرارات :

حالات الاختبار	$\alpha: \alpha$ ترفض H_0 اذا كانت قيمة	$\alpha: 0.05$ ترفض H_0 اذا كانت قيمة	$-\alpha: 0.01$ ترفض H_0 اذا كانت قيمة
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$	$Z \geq Z_{\frac{\alpha}{2}}$ and $Z \leq -Z_{\frac{\alpha}{2}}$ or $ Z \geq Z_{\frac{\alpha}{2}}$	$Z \geq 1.96$ and $Z \leq -1.96$ or $ Z \geq 1.96$	$Z \geq 2.58$ and $Z \leq -2.58$ or $ Z \geq 2.58$
$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$	$Z \geq Z_{\alpha}$	$Z \geq 1.65$	$Z \geq 2.33$
$H_0: \mu = \mu_0$	$Z \leq -Z_{\alpha}$	$Z \leq -1.65$	$Z \leq -2.33$

H1:	$\mu < \mu_0$		
-----	---------------	--	--

مثال:

ينتج معمل للتعليب قناني فاكهة مفروض ان يكون متوسط وزنها (15) باوند وبانحراف قياسي هو (0.5) باوند للتاكيد من ان المعمل لازال ينتج عند المستوى المطلوب اخذت عينة مكونة من (50) علبة فوجد ان متوسط وزنها (14.8) باوند فاذا كان وزن العلبة متغير عشوائي تتوزع توزيع طبيعيًا فهل تدل العينة على ان انتاج المعمل لازال (15) باوند باختبار $\alpha = 0.01$ كمستوى معنوية؟

الحل :

-فرضية العدم $H_0: \mu = 15$

-الفرضية البديلة $H_1: \mu \neq 15$

-مستوى المعنوية $\alpha = 0.01$

-منطقة الرفض $Z \geq 2.58$ or $Z \leq -2.58$

-المختبر الاحصائي $x=14.8, n=50, \sigma=0.5$

$$Z = \frac{x - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{14.8 - 15}{\frac{0.5}{\sqrt{50}}} = -2.83$$

-القرار : بما ان قيمة Z المحسوبة (-2.83) هي اقل من قيمة Z الجدولية (-2.58) أي ان قيمة Z المحسوبة تقع في منطقة الرفض لذا ترفض فرضية العدم وتقبل الفرضية البديلة

$H_0: \mu = 15$

$H_1: \mu \neq 15$

أي ان المعمل لا ينتج علب اوزانها (15) باوند

مثال :

كان معدل انتاج احد الاصناف المحلية من الحنطة في السنين الخمس السابقة هو 1600 كغم / هكتار وقد ادعى احد المزارعين بانه قد استنبط سلالة من هذا الصنف يعطي انتاج اكثر

ولاختبار صحة الادعاء اخذت عينة عشوائية مولفة 81 من نباتا من السلالة الجديدة ووجد ان متوسط انتاجها 1630كغم / هكتار بانحراف قياسي 15 كغم هل نتائج العينة تؤيد ادعاء الباحث تحت مستوى معنوي 0.01 ؟

الحل: فرضيةالعدم $H_0: \mu = 1600$

-الفرضية البديلة $H_1: \mu > 1600$

مستوى المعنوية $\alpha = 0.01$

منطقة الرفض $Z \geq 2.33$

-المختبر الاحصائي $x=1630, n = 81, \sigma = 15$

$$Z = \frac{x - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1630 - 1600}{\frac{15}{\sqrt{81}}} = 18$$

-القرار : بما ان قيمة Z المحسوبة (18) هي اقل من قيمة Z الجدولية (2.33) أي ان قيمة Z المحسوبة تقع في منطقة الرفض لذا ترفض فرضية العدم وتقبل الفرضية البديلة أي ان ادعاء المزارع كان صحيحا .

مثال /

ادعت احدى شركات انتاج البنجر السكري بانها انتجت صنفا من البنجر السكري نسبة السكر فيه لا تقل عن (0.18) غم/ أي 18لكل (100)غم وبانحراف قياسي قدره (2.5) غم ولاختبار هذا الاعاء اخذت عينة عشوائية مؤلفة من (36) ثمرة من البنجر وحسبت منه نسبة السكر فكان وسطها الحسابي(17.2)غم/(100)غم فهل ادعاء الشركة عند مستوى احتمال (0.05) مقبولا؟
الحل:

فرضيةالعدم $H_0: \mu = 18$

-الفرضية البديلة $H_1: \mu < 18$

مستوى المعنوية $\alpha = 0.05$

منطقة الرفض $Z \leq -1.65$

-المختبر الاحصائي $x=17.2, n = 36, \sigma = 2.5$

$$Z = \frac{x - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{17.2 - 18}{\frac{2.5}{\sqrt{36}}} = -1.9$$

القرار : بما ان قيمة Z المحسوبة (-1.9) هي اقل من قيمة Z الجدولية (-1.65) أي ان قيمة Z المحسوبة تقع في منطقة الرفض لذا ترفض فرضية العدم وتقبل الفرضية البديلة أي بعبارة اخرى ان ادعاء الشركة غير صحيح تحت مستوى احتمال 0.05 .

واجب /

نفس المثال ولكن استخدم $\alpha = 0.01$

اختبارات تتعلق بمتوسطين Tests Concerning Two Means .

في بعض الاحيان نحتاج مقارنة متوسطي مجتمعين ، فمثلا نقارن بين نوعين من التغذية للاطفال او مقارنة صنفين من الذرة الصفراء.....الخ

ففي هذه الحالة سنتضمن الفرضية مقارنة الفرق بين متوسطين هما μ_1 ، μ_2 وتباينهما θ_1 ، θ_2 وسوف نختبر الفرضية القائلة ان الفرق بين المتوسطين يساوي قيمة معينة أي

$$H_0: \mu_1 - \mu_2 = d_0$$

حيث ان هي القيمة المعلومة ، اما الفرضية البديلة ستكون احدى الفرضيات الاتية

$$H_0: \mu_1 - \mu_2 = d_0$$

$$H_1: \mu_1 - \mu_2 > d_0$$

$$H_1: \mu_1 - \mu_2 < d_0$$

مفهوم درجة الحرية Concept of degree of freedom

لتوضيح هذا المفهوم نفرض ان عينة حجمها (5) أي (n=5) وكما هو معلوم فان مجموع الانحرافات عن الوسط الحسابي يساوي صفر أي ان

$$\sum_{i=1}^n (x_i - \bar{x}) = 0$$

لذا سيكون لدينا (5) انحرافات ولكن d_1, d_2, d_3, d_4, d_5 فاذا عرفنا اربعة من هذه الانحرافات فاننا بسهولة سنعرف الخامس فليس لنا حرية الاختيار قيمة له لانه قد حددت قيمة بعد معرفة القيم الاربعة مثلا :

$$d_1 = 5 , d_2 = -3 , d_3 = 1, d_4 = 6$$

والمطلوب ايجاد قيمة d_5 هي

$$d1 + d2 + d3 + d4 + d5 = 0$$

$$5 + (-3) + 1 + 6 + d5 = 0$$

$$D5 = 9$$

لذا فاذا كان لدينا عينة حجمها n فلنا الحرية في الاختيار قدرها $(n-1)$ لاختيار قيم الانحرافات عن الوسط الحسابي للعينة لذا فان مجموع مربعات الانحرافات عن الوسط الحسابي لها درجات الحرية قدرها $(n-1)$ وبالتالي فان درجات الحرية للتباين هو ايضا $(n-1)$

مثال:

لوحظ من خلال الفترة السابقة ان متوسط المبيعات الاسبوعية من سلعة معينة كان (150) صندوق قامت الشركة المنتجة لهذه السلعة بعمل اعلان تلفزيوني الهدف منه رفع متوسط المبيعات الاسبوعية من تلك السلعة وبعد مضي فترة زمنية لوحظ ان متوسط المبيعات من هذه السلعة في (24) محل كان (172) صندوق والانحراف القياسي (27.4) صندوق هل ترى لهذا الاعلان اثر جوهري في رفع متوسط المبيعات الاسبوعية عند مستوى معنوي (0.05) ؟
الحل/

$$n = 24 , S = 27.4 , x = 172 , u = 150$$

صياغة الفروض

$$H_0 : \mu = 150$$

$$H_1 : \mu > 150$$

$$\alpha = 0.05$$

مستوى المعنوية

$$T(23, 1-0.05) = t(23, 0.95) = 1.714$$

حساب احصائية الاختبار

$$t^* = \frac{x - \mu}{\frac{s}{\sqrt{n}}} = \frac{172 - 150}{\frac{27.4}{\sqrt{24}}} = \frac{22}{5.74} = 3.933$$

$$= 3.933$$

اتخاذ القرار :

قيمة المحسوبة اكثر من قيمة t الجدولية عند درجة حرية (23) ومستوى المعنوية 0.05 التي هي (1.714) لذا نرفض H_0 ونقبل H_1 أي ان الاعلان له اثر جوهري في رفع متوسط المبيعات الاسبوعية .

الفصل السادس

تحليل التباين Analysis of Variance

تحليل التباين بمعيار واحد One way analysis of Variance

من المثال التالي نشرح طريقة الحل :

لصناعة منتج معين استخدمت مواد اولية من ثلاثة مصادر مختلفة اختيرت عينات من القطع المنتجة لكل مادة اولية مستعملة بقصد معرفة ما اذا كان هناك فرق في نسبة الشوائب من القطع المنتجة من اختلاف مصدر المواد المستخدمة تحت مستوى اختبار 0.05 علما بان منطقة الرفض (8,02)

المصدر 1	1	2	2	4
المصدر 2	3	4	5	7
المصدر 3	6	8	8	9

الحل:

-نحسب مجموع المفردات ومجموع مربعات المفردات للعينات

$$T1=1+2+2+4=9 \quad Ts1=1+4+4+16=25$$

$$T2=3+4+5+7 =19 \quad Ts2=9+16+25+49=99$$

$$T3=6+8+8+9=31 \quad Ts3=36+64+64+81=245$$

-نحسب المجموع الكلي للمفردات والمجموع الكلي لمربعات المفردات:

$$T = \sum_{i=1}^3 T_i = 59$$

$$Ts = \sum_{i=1}^3 T_{si} = 369$$

-نحسب مجموع المربعات الكلي (Total sum of squares) Tss

$$Tss = Ts - T^2 / N = 369 - 290.1 = 78.9$$

-نحسب مجموع المربعات الاختلافات بين الفئات (SSB)

Sum of Squares Between Classes

$$SSB = \frac{T1^2}{n1} + \frac{T2^2}{n2} + \frac{T3^2}{n3} - \frac{T^2}{N}$$

$$SSB = \frac{9^2}{4} + \frac{19^2}{4} + \frac{31^2}{4} - \frac{(59)^2}{12} =$$

$$\frac{81}{4} + \frac{361}{4} + \frac{961}{4} - \frac{3481}{12} =$$

$$20.25 + 90.25 + 240.25 - 290.1 = 60.65$$

- نحسب المجموع الكلي للمربعات داخل الفئات (SSW)

Sum of Squares Within Classes

وغالبا ما يصطلح عليها بالخطأ العشوائي

$$SSW = Tss - SSB = 78.9 - 60.65 = 18.25$$

وعليه سيكون جدول تحليل التباين :

مصدر التباين	درجة الحرية (df)	مجموع المربعات (ss)	متوسط المربعات (μ_s)
بين المجموعات	3-1	SSB = 60.65	60.65 / 2 = 30.33
داخل المجموعات	12 - 3 = 9	SSW = 18.25	18.25 / 9 = 2.03
المجموع	11	78.9	14.940

عند مستوى معنوي (0.05) نستخرج من جدول توزيع F القيمة المرادفة تحت درجات

$$F_{2,9} = 8.02 \quad \text{حرية (9,2)}$$

وبما ان قيمة F المحسوبة اكبر من قيمة F الجدولية بدرجة حرية (2,9) نستنتج من اخيبارنا هذا بان نسبة الشوائب في قطع الانتاج تتاثر باختلاف المادة الاولية المستعملة أي اننا رفضنا الفرضية الاحصائية بعدم وجود فرق معنوي في نسبة الشوائب علما بانه نسبة الخطأ في استنتاجنا هذا هي 5%

مثال :

: زرعت ثمانية انواع مختلفة من الحنطة في ثلاثة قطع اراضي فاعطتنا النتائج الموجودة في الجدول ادناه اختبر عند مستوى معنوية 0.05 فيما اذا كان معدل الانتاج مساويا للانواع الثمانية من الحنطة علما بان منطقة الرفض

T1	T2	T3	T4	T5	T6	T7	T8	المجموع
8	5	8	10	5	12	5	7	

7	4	7	9	4	8	7	8	
9	7	9	6	11	6	10	8	
24	16	24	25	20	26	22	23	180

نلاحظ في هذا المثال ان عدد الاعمدة (k=8) وعدد الصفوف (n=3)

نحسب مجموع المفردات ومجموع مربعات المفردات للعينات

$$T1=8+7+9=24$$

$$Ts1=64+49+81=194$$

$$T2=3+4+5+7 =19$$

$$Ts2=9+16+25+49=99$$

$$T3=8+7+9=24$$

$$Ts3=64+49+81=194$$

$$T4 =10+9+6= 25$$

$$Ts4=100+81+36=217$$

$$T5=5+4+11=20$$

$$Ts5=25+16+121=217$$

$$T6=12+8+6= 26$$

$$Ts6 =144+64+36= 244$$

$$T7=5+7+10= 22$$

$$Ts7= 25+49+100= 174$$

$$T8= 7+8+8= 23$$

$$Ts8= 49+64+64= 177$$

نحسب المجموع الكلي للمفردات والمجموع الكلي لمربعات المفردات:

$$T = \sum_{i=1}^8 Ti = 180$$

$$Ts = \sum_{i=1}^8 Tsi = 1452$$

-نحسب مجموع المربعات الكلي (Total sum of squares) Tss

مجموع مربعات انحرافات المفردات عن متوسط الحسابي العام للمجتمع ويمكن حسابه

بالطريقة الاتية

$$Tss = Ts - T^2 / N = 1452 - (180)^2 / 24 =$$

$$1452 - 1350 = 102$$

-نحسب مجموع المربعات الاختلافات بين الفئات (SSB)

Sum of Squares Between Classes

$$SSB = \frac{T1^2}{n1} + \frac{T2^2}{n2} + \frac{T3^2}{n3} + \frac{T4^2}{n4} + \frac{T5^2}{n5} + \frac{T6^2}{n6} + \frac{T7^2}{n7} + \frac{T8^2}{n8} - \frac{T^2}{N}$$

$$SSB = \frac{24^2}{3} + \frac{16^2}{3} + \frac{24^2}{3} + \frac{25^2}{3} + \frac{20^2}{3} + \frac{26^2}{3} + \frac{22^2}{3} + \frac{23^2}{3} - \frac{(180)^2}{24}$$

$$= 192+85.3+192+208.3+133.3+225.3+161.3+176.3-1350$$

$$= 1374 - 1350 = 24$$

- نحسب المجموع الكلي للمربعات داخل الفئات (SSW)

Sum of Squares Within Classes

وغالبا ما يصطلح عليها بالخطأ العشوائي

$$SSW = Tss - SSB = 102 - 24 = 78$$

درجات الحرية للاعمدة $v1=k-1=7$, $v2=k(n-1)=16$ وداخل المجموعات (للخطأ)

$$F = \frac{\frac{SSB}{k-1}}{\frac{SSW}{k(n-1)}}$$

$$= \frac{\frac{24}{7}}{\frac{78}{16}} = \frac{3.429}{4.875} = 0.703$$

من جدول توزيع f وعند مستوى معنوية (0.05) ودرجات الحرية $v1=7, v2=16$

نجد ان $f_{7,16} = 2.66$ تمثل منطقة الرفض وبما ان (0.703) لا تقع في منطقة الرفض اذن تقبل H_0 .

الفصل السابع:

الانحدار والارتباط البسيط

Simple Regression and Correlation

مقدمة :

في الفصول السابقة كان اهتمامنا ينصب على قضايا الاحصاء الاستنتاجي التي تعود الى توزيع ذو متغير واحد. هو X

اما الان فسنحول اهتمامنا الى قضايا تخص توزيع ذو متغيرين

(Birariate Dist) وسنرمز لهذين المتغيرين بالرمز X, Y فمثلا قد يكون المتغير X هو عدد نبات القطن في وحدة المساحة بينما المتغير Y هو كمية المحصول الناتج او قد يكون X هو درجات الحرارة بينما المتغير Y هو الكمية المذابة غم من الماء في مادة كيميائية معينة او قد تكون X هو المعدل الفصلي للطلبة بينما Y هو الدرجات النهائية لهم في مادة الاحصاء ومن ذلك يتضح بان كل فرد من الافراد العينة له قياسان احدهما للمتغير X والآخر للمتغير Y فمثلا لكل طالب درجتان هما معدله الفصلي (X) ودرجته النهائية (Y)

وهناك نوعان من الانحدار :

- انحدار بسيط (Simple Regression)

- انحدار متعدد (Multiple Regression)

الانحدار البسيط: (Simple Regression)

يشمل على متغيرين فقط هما مستقل (X) و معتمد (Y)

اما الانحدار المتعدد: (Multiple Regression)

يشمل على اكثر من متغيرين احدهما معتمد (Y) و مستقل (x_1, x_2, \dots, x_n)

مثال:

البيانات التالية تمثل الدرجة الفصلية و الدرجة النعائية في درس الاحصاء لاثني عشر طالبا . اوجد معادلة الانحدار ؟

X_i	65	50	55	65	55	70	65	70	55	70	50	55
Y_i	85	74	76	90	85	87	94	98	81	91	76	74

الحل :

X_i	65	50	55	65	55	70	65	70	55	70	50	55
Y_i	85	74	76	90	85	87	94	98	81	91	76	74
X_i^2	4225	2500	3025	4225	3025	7569	4225	7569	3025	7569	2500	3025
$X_i Y_i$	5525	3700	4180	5850	4675	6090	6110	6860	4455	6370	3800	4070

$$b = \frac{\sum x_i y_i - [(\sum x_i)(\sum y_i)]/n}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}$$

$$= \frac{61685 - \frac{(725)(1011)}{12}}{4475 - \frac{(725)^2}{12}} = 0.897$$

$$a = y - b x$$

$$a = 84.250 - (0.897)(60.417) = 30.056$$

$$Y^* = Y = a + b x = 30.056 + (0.897)x$$

أي قيمة نأخذ ل X ولتكونان 50 , 70 سوف نجد منها قيمة ال Y^* و Y فتكون:

$$Y(50) = 30.056 + 0.897(50) = 74.9$$

$$Y(70) = 30.056 + 0.897(70) = 92.8$$

والخط المستقيم الذي يمر بها تين النقطتين هو الانحدار

وبالتعويض في هذه المعادلة عن كل قيمة من قيم فانه يمكن حساب قيم تقديرية ال (Y, Y^*) والتي تقع جميعها على خط الانحدار البسيط وهذه القيم موضحة في الجدول التالي

X_i	65	50	55	65	55	70	65	70	55	70	50	55
Y_i	85	74	76	90	85	87	94	98	81	91	76	74
$Y(x)$	88.4	74.9	79.4	88.4	79.4	92.8	88.4	92.8	79.4	92.8	74.9	79.4

اذن خطوات الحل لايجاد معادلة لانحدار الخطي هي :

- نجد X_i^2 و $X_i Y_i$ في الجدول

- ثم نجد a, b

- ثم نجد معادلة خط الانحدار والتي هي:

$$Y^* = Y(x) = a + b x$$

ومن خواص معادلة الانحدار الخطي البسيط هو

- انقيمة النقطة (x_i, y_i) تقع على خط الانحدار

- ان خط الانحدار يمر من جميع قيم (x_i, y_i)

- ان مجموع الانحرافات عن خط الانحدار يساوي صفر أي ان $\sum(y_i - yx) = 0$

- ان مجموع مربعات الانحرافات عن خط الانحدار هي اقل ما يمكن أي ان

$$\sum(y_i - yx)^2 = \text{minmum}$$

هذا يمكن ان نستخدم معادلة خط الانحدار للتنبؤ عن قيمة Y لقيمة معينة من Y غير

موجودة في العينة مثل الدرجة النهائية المتوقعة لطالب معدله الفصلي (73) هي

$$Y(73) = 30.056 + 0.897 (73) = 95.5$$