8-2-4 توزيع χ^2 (مربع كاي)

ان توزيع مربع كاي من التوزيعات الاحتمالية المستمرة والمهمة في اختبار الفرضيات علما ان توزيع χ^2 هو في الحقيقة حالة خاصة من توزيع كاما Γ .

ان دالة كثافة الاحتمال للمتغير العشوائي χ^2 هي

$$f(\chi^2) = K(\chi^2)^{\frac{v-2}{2}} e^{-\frac{\chi^2}{2}} \qquad \chi^2 > 0$$

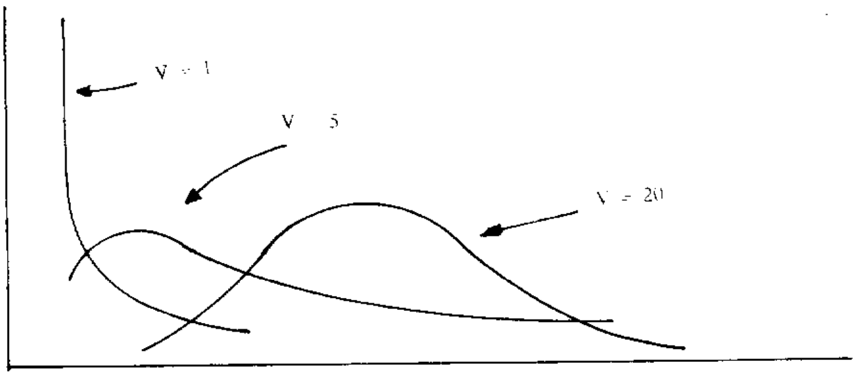
حيث ان :

$e = 2.71828$

$v =$ درجات الحرية

$$K = \frac{1}{\frac{(v/2)}{2}\,\Gamma\left(\frac{v}{2}\right)}$$

اما المتوسط الحسابي لتوزيع χ^2 فانه يساوي (V) .
والتباين يساوي (2V) .
وعند معرفتنا لدرجة الحرية فمن السهولة تحديد توزيع χ^2 كما في الشكل 4–6 :



شكل (4 – 6) توزيع χ^2 لعدة درجات حرية

ان لتوزيع مربع كاي عدة خواص رياضية مهمة ندرجها كما يلي :

١ . اذا كان Y يتوزع توزيعا معتدلا قياسيا بمتوسط حسابي صفر وتباين = ١، فان Y^2 له توزيع χ^2 بدرجة حرية = ١

٢ . اذا كان كل من

$$Y_1 , Y_2 \ldots Y_n$$

متغيرات عشوائية مستقلة لها توزيع معتدل قياسي فان :

$$\Sigma Y_i^2$$

له توزيع مربع كاي بدرجة حرية n

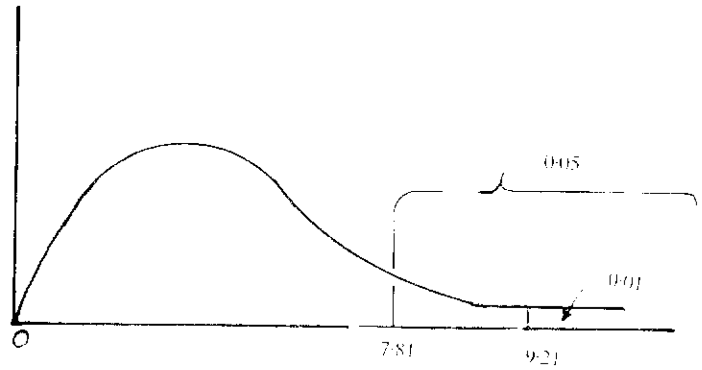
٣ . $\chi^2 (n \pm m) = \chi^2 n \pm \chi^2 m$

٤ . اذا كان S^2 هو تباين عينة عشوائية ذات حجم n اخذت من مجتمع معتدل له متوسط حسابي μ وتباين σ^2 . فان

$$\frac{(n - 1) S^2}{\sigma^2}$$

له توزيع χ^2 بدرجات حرية مساوية الى $(n - 1)$.

وقد قام باحثون بحساب قيم χ^2 لدرجات حرية متفاوتة . ويمثل الجدول الملحق رقم (5) بعض هذه القيم التي تحدد مساحات مساوية الى 0.05 و 0.01 عند نهاية التوزيع . ويمثل الشكل جـ 7 طريقة استخدام هذا الجدول بالنسبة لدرجات حرية مساوية الى 3 .



شكل (4 –) : تحديد القيم الجدولية لتوزيع χ^2

3-4 العلاقة بين احصاءات العينة وثوابت أو معالم المجتمع

لقد سبق وان اشرنا الى ان الباحث عادة مايضطر لاخذ عينة من مجتمع معين بهدف تعميم النتائج التي تسفر عنها بيانات العينة على المجتمع الذي اخذت منه شريطة كون تلك العينة ممثلة لذلك المجتمع. وتعتبر تقديرات معالم او ثوابت المجتمع (مثل المتوسط الحسابي والانحراف المعياري والنسبة) المستندة على نتائج العينات نوعا من هذا التعميم. وتستند مثل هذه التقديرات على علاقات محددة بين معالم المجتمع واحصاءات العينة المناظرة. وسنركز في بحثنا على العلاقة القائمة بين المتوسط الحسابي ($\bar{X}$) التباين (S_x^2) والنسبة ($\bar{P}$) للعينة ومعالم المجتمع المناظرة لها اي (P , σ_x^2 , μ_x)

1-3-4 علاقة المتوسط الحسابي للعينة بالمتوسط الحسابي للمجتمع :

لتوضيح العلاقة القائمة بين المتوسط الحسابي للعينة ($\bar{X}$) والمتوسط الحسابي للمجتمع (μ_x) للمتغير (X) فاننا سنستعين بمثال يجعل ادراك هذه العلاقة امرا سهلا ومفيدا.

مثال 4-6 لو كان لدينا مجتمعا مكونا من خمسة رؤوس من الغنم وان اوزان هذه الرؤوس هي :

رمز الرأس	الوزن (كغم) (المتغير X)
أ	27
ب	39
ج	30
د	36
ه	42

فلو اطلقنا على الوزن بالمتغير (X) . فان المتوسط الحسابي لهذا المجتمع هو :

$$\mu_x = \frac{\Sigma X i}{N}$$

$$= \frac{27 + 39 + 30 + 36 + 42}{5}$$

$$= 34{\cdot}8 \text{ kg}$$

والان . دعنا نفترض أننا . لسبب أو لاخر . نرغب في أخذ عينة عشوائية ممثلة لهذا المجتمع ومكونة من ثلاثة رؤوس بهدف تعميم نتائجها على المجتمع نفسه . ولا يخفى في هذه الحالة ان نتائج العينة تعتمد على المفردات (الرؤوس في مثالنا الحالي) التي تشملها العينة . وبما أن لدينا خمسة رؤوس . فان العينة المكونة من ثلاثة رؤوس يمكن ان تكون أي عينة من العينات العشرة التالية الممكن تشكيلها من الخمسة رؤوس :
أ ب جـ أ ب د أ ب هـ أ جـ د أ جـ هـ أ د هـ ب جـ د ب جـ هـ
ب د هـ جـ د هـ .

ولا بد لنا من الاشارة الى أن ترتيب **الحروف** لأي عينة كانت تعطينا نفس تلك العينة للاغراض التي **نحن** بصددها . أي ان :

أ ب جـ أ جـ ب ب أ جـ ب جـ أ جـ أ ب جـ ب أ هي نفس **العينة** لانها جميعا تشتمل على الرؤوس أ و ب و جـ .

واذا ماعدنا الى العينات العشرة الممكن تشكيلها من مجتمعنا هذا . فان اعتمادنا على العينة المختارة في اعطاء فكرة عن معالم المجتمع (المتوسط الحسابي في هذه الحالة) يعتمد على نتائج تلك العينة . ولا يخفى ان النتائج التي نحصل عليها من العينات العشرة لا تتشابه بالضرورة . وفي الواقع . فاننا سنحصل على متوسطات حسابية مختلفة للعينات العشرة كما هو موضح في جدول 4–2

جدول 4-2 العينات العشرة الممكنة ومتوسطاتها الحسابية

رقم العينة	مفردات العينة	أوزان مفردات العينـــة (كغـــم)	المتوسط الحسابـي للعينة (X) كغـــم
1	أ . ب . جـ	30,39,27	32
2	أ . ب . د	36,39,27	34
3	أ . ب . هـ	42,39,27	36
4	أ . ب . د	36,30,27	31
5	ا . جـ . د	42,30,27	33
6	أ . د . هـ	42,36,27	35
7	ب . جـ . د	36,30,39	35
8	ب . جـ . هـ	42,30,39	37
9	ب . د . هـ	42,36,39	39
10	جـ . د . هـ	42,36,30	36

ومن تفحص المتوسطات الحسابية للعينات العشرة يبدو واضحا أنها تختلف من عينة الى أخرى . كما أن أي منها لا تساوي المتوسط الحسابي للمجتمع الذي أخذت منه والبالغ 34.8 كغم . ونلاحظ كذلك ان بعض المتوسطات الحسابية للعينات العشرة أقل من المتوسط الحسابي للمجتمع والبعض الاخر اكبر منه . ان المتوسطـات الحسابيـة العشرة تشكل مجتمعا قائما بذاته لانها القيم الوحيدة الممكن الحصول عليها من عينات حجم الواحدة منها ثلاثة رؤوس من المجتمع المكون من خمسة رؤوس . وعليه فان بامكاننا الان أن نتكلم عن مجتمع محدد يمثل متوسطات حسابية للعينات العشرة التي تختلف من عينة الى أخرى . أي أن المتغير الذي نحن بصدده هو المتوسط الحسابي للعينـة (أي $\bar{X}$) . فاذا أردنا حساب المتوسط الحسابي لهذا المجتمع نحصل علـى ما يلـي :

$$\mu_x = \frac{\Sigma X_i}{N}$$

$$= \frac{32 + 34 + 36 + 31 + 33 + 35 + 35 + 37 + 39 + 39}{10}$$

$$= \frac{348}{10}$$

$$= 34.8 \text{ Kg}.$$

وهذه الحقيقة تعني أن :

$$\mu_{\bar{x}} = \mu_x \qquad \ldots (8 - 4)$$

أي أن المتوسط الحسابي للمتوسطات الحسابية لجميع العينات الممكن أخذها من مجتمع معين يساوي بالضبط المتوسط الحسابي للمجتمع نفسه . ولهذا السبب يعتبر المتوسط الحسابي للعينة ($\bar{X}$) تقديرا غير متحيز للمتوسط الحسابي للمجتمع (μ_x) .

4-3-2 علاقة الانحراف المعياري للعينة بالانحراف المعياري للمجتمع

يمكننا باسلوب مماثل ، أن نبين بان تباين العينة (S_x^2) المحسوب وفق المعادلة التاليـة :

$$S_x^2 = \frac{\Sigma (X - \bar{X})^2}{n - 1}$$

هو تقدير غير متحيز لتباين المجتمع (σ_x^2) اذا كان المجتمع المأخوذة منه العينة هو مجتمع غير محدود الحجم . وهذه الحقيقة تجعل من الصعب اختيار مثال عملي لتوضيح هذه الفكرة . وعليه ، فاننا سنكتفي بتدوين العلاقة القائمة بين تباين العينـة (S_x^2) وتباين المجتمع (σ_x^2) على النحو المبين في المعادلة (4 - 9) :

$$\mu_{S_x^2} = E(S_x^2) = \sigma_x^2 \qquad \ldots (9-4)$$

أي ان المتوسط الحسابي لتباين جميع العينات ذات حجم (n) الممكن أخذها من مجتمع كبير جدا أو لا محدود يساوي تباين المجتمع نفسه من الناحية النظرية على الاقل .

4-3-3 علاقة نسبة العينة بنسبة المجتمع

اذا اردنا ان نعطي فكرة عن نسبة الاصابة في مجتمع من النباتات أو الحيوانات فان بالامكان حساب هذه النسبة بشمول جميع أفراد المجتمع وتصنيفها الى مصاب وغير مصاب أو أخذ عينة ممثلة له وحساب نسبة الاصابة بين افرادها واعتبارها تقديرا لنسبة الاصابة في المجتمع . ولتمييز نسبة العينة عن نسبة المجتمع باعتبار الأولى من بين الاحصاءات والثانية من الثوابت أو المعالم فاننا نرمز لنسبة العينة بالرمز ($\bar{P}$) ونسبة المجتمع بالرمز (P) .

ولمعرفة العلاقة بين النسبتين ، يمكننا الاستعانة بالمثال 4 - 6 . فلو فرضنا ان النسبة التي نحن بصددها هي نسبة الاغنام التي يقل وزنها عن 35 كغم .

وبما أن هناك رأسان فقط تقل اوزانهما عن 35 كغم . فان نسبة الأوزان التي تقل عن 35 كغم في المجتمع هي :

$$P = \frac{2}{5} = 0.4$$

أما نسبة الأوزان التي تقل عن 35 كغم في العينة المكونة من ثلاثة رؤوس فانها تختلف من عينة الى عينة كما هو مبين في الجدول رقم 4 - 3 :

جدول 4 - 3 نسبة الأوزان التي تقل عن 35 كغم في العينات العشرة

رقم العينة	مفردات العينة	اوزان مفردات العينة	نسبة الأوزان التي تقل عن 35 كغم في العينة ($\bar{P}$)
1	أ ، ب ، جـ	30 ، 39 ، 27	2/3
2	أ ، ب ، د	36 ، 39 ، 27	1/3
3	أ ، ب ، هـ	42 ، 39 ، 27	1/3
4	أ ، جـ ، د	36 ، 30 ، 27	2/3
5	أ ، جـ ، هـ	42 ، 30 ، 27	2/3
6	أ ، د ، هـ	42 ، 36 ، 27	1/3
7	ب ، جـ ، د	36 ، 30 ، 39	1/3
8	ب ، جـ ، هـ	42 ، 30 ، 39	1/3
9	ب ، د ، هـ	42 ، 36 ، 39	0
10	جـ ، د ، هـ	42 ، 36 ، 30	1/3

وبما أن نسب العينات ($\bar{p}$) تشكل مجتمعا قائما بذاته ، فان متوسطها الحسابي هـو :

$$\mu_{\bar{p}} = \frac{(3)(\frac{2}{3}) + 6(\frac{1}{3}) + (1)(0)}{10}$$

$$= \frac{4}{10}$$

$$= 0.4$$

ويلاحظ ان المتوسط الحسابي لمجتمع نسب العينات الممكن أخذها من مجتمع معين يساوي بالضبط لنسبة المجتمع . أي أن :

$$\mu_{\bar{p}} = p \qquad \ldots (10-4)$$

وهذا يشكل دليلا على ان نسبة العينة ($\bar{P}$) تعتبر **تقديراً** غير متحيز لنسبة المجتمـــع .

وتجدر الاشارة الى ان **سبب** اختيار الرمز ($\bar{P}$) لنسبة العينة هو لان هذه النسبة ماهي الا متوسط حسابي لتصنيف المفردات الى اقل من 35 كغم واكثر من 35 كغم والاستعاضة بالمفردة التي يقل وزنها عن 35 كغم بالقيمة الرمزية (١) ولتلك التي يزيد وزنها عن كغم بالقيمة الرمزية (0) وجمع هذه القيم وتقسيمها على عددها كما هو متبع فـــي تحديد المتوسط الحسابي . فلو اخذنا العينة رقم (١) في مثالنا الحالي فان تصنيــف المفردات الى اقل واكثر من 35 كغم وقيمها الرمزية يأخذ الشكل التالي :

مفردات العينة	القيم الرمزية
أ . ب . ج	١ . 0 . ١

وعليه فان نسبة الرؤوس ذات الوزن الاقل من 35 كغم هي :

$$\bar{P} = \frac{1+0+1}{3} = 2/3$$

4-3-4 علاقة الانحراف المعياري للمتغير ($\bar{X}$) بالانحراف المعياري للمجتمع :

يمكن تحديد العلاقة بين الانحراف المعياري للمتوسط الحسابي للعينة (اي $\sigma_{\bar{x}}$ والانحراف المعياري للمجتمع (σ_x) من العودة الى المثال 4 - 6 . وبما ان المتوسطات الحسابية العشرة التي تم التوصل اليها في المثال المذكور تشكل مجتمعا قائما بذاته . فان الانحراف المعياري للمتغير ($\bar{X}$) والذي يطلق عليه (الخطأ المعياري Standard Error) تمييزا له عن الانحراف المعياري للمشاهدة الواحدة . يمكن حسابه على النحو التالي :

$$\sigma_{\bar{x}} = \sqrt{\sigma_{\bar{x}}^2} = \sqrt{\frac{\Sigma \bar{X}_i^2 - \frac{(\Sigma \bar{X}i)^2}{N}}{N}}$$

حيث ان N تمثل حجم مجتمع المتوسطات الحسابية .

$$= \sqrt{\frac{12162 - \frac{(348)^2}{10}}{10}}$$

$$= \sqrt{5.16}$$

$$= 2.3 \text{ Kg}.$$

وهذه الطريقة تعني ان حساب الخطأ المعياري ($\sigma_{\bar{x}}$) يتطلب تحديد المتوسط الحسابي لكل من العينات الممكن اخذها من المجتمع وبما ان هذا الاسلوب غير ممكن من الناحية العملية وفي أغلب الاحيان وخاصة عندما يكون حجم المجتمع كبيرا فان العلاقة الرياضية التالية تسمح بحساب الخطا المعياري بصورة مباشرة ومن المعادلة (4 - 11) :

$$\sigma_{\bar{x}} = \sqrt{\frac{N - n}{N - 1}} \cdot \frac{\sigma_x}{\sqrt{n}} \qquad \ldots (11 - 4)$$

حيث أن :

N يمثل حجم المجتمع

n يمثل حجم العينة

σ_x يمثل الانحراف المعياري للمجتمع

وعليه فان المعادلة (4 – 11) تمكننا من حساب الخطأ المعياري لمجرد معرفة N ، n ، σ_x وتغنينا عن حصر جميع العينات الممكن أخذها من المجتمع ومايترتب على ذلك من عمليات حسابية مرهقة اضافة الى استحالة امكانية اجراء ذلك في معظم الحالات .

فاذا طبقنا المعادلة (4 - 11) على مثالنا اعلاه والذي يتميز بالصفات التاليـــة :

$$N = 5$$
$$n = 3$$

$$\sigma_x = \sqrt{\frac{\Sigma X_i^2 - \frac{(\Sigma X i)^2}{N}}{N}}$$

$$= \sqrt{30{\cdot}96}$$

$$= 5{\cdot}6 \text{ Kg}.$$

وعوضنا هذه القيم في المعادلة (4 11) نحصل على مايلي :

$$\sigma_{\bar{x}} = \sqrt{\frac{5-3}{5-1}} \cdot \frac{5{\cdot}6}{\sqrt{3}}$$

$$= 2{\cdot}3 \text{ kg}$$

ويلاحظ باننا حصلنا على نفس القيمة التي توصلنا اليها بواسطة الطريقة المطولـــــة .

وهناك معادلات ذات علاقة بالمعادلة (4 - 11) وهي :

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} \qquad \dots (12 - 8)$$

في حالة كون المجتمع غير محدود الحجم أو أن حجمه كبير جداً بحيث تؤول السكمية $(N-n)/(N-1)$ الى الواحد الصحيح . وتجدر الاشارة الى ان هذه السكمية يطلق عليها « معامل التصحيح للمجتمع محدود الحجم » .

وبما ان الانحراف المعياري للمجتمع (σ_x) غيرمعلوم في الحالات الاعتيادية ، فان بالامكان الاستعاضة عنه بالانحراف المعياري للعينة (S_x) . اعتماداً على المعادلة (4 – 9) التي تشير الى ان (S_x) هوتقدير غيرمتحيز (σ_x) وعليه فان :

$$S_{\bar{x}} = \sqrt{\frac{N-n}{N-1}} \cdot \frac{S_x}{\sqrt{n}} \qquad \ldots (13-4)$$

وذلك في حالة عدم معرفة (σ_x) مع معرفة حجم المجتمع (N) وأما

$$S_{\bar{x}} = \frac{S_x}{\sqrt{n}} \qquad \ldots (14-4)$$

في حالة المجتمع اللامحدود او غيرمعروف الحجم . ولايخفى ان هذه الحالة هي أكثر الحالات شيوعاً في الحياة العملية .

4 – 3 – 5 الانحراف المعياري للنسبة في العينة

اذا أتبعنا نفس الاسلوب المشار اليه في البند السابق فأننا نستطيع تحديد الأنحراف المعياري للنسبة في العينة ($\sigma_{\bar{p}}$) في حالة المجتمع محدود الحجم وفق المعادلة

$$\sigma_{\bar{x}} = \sqrt{\frac{N-n}{N-1}} \cdot \sqrt{\frac{P(1-P)}{n}} \qquad \ldots (15-4)$$

أما في حالة المجتمع اللامحدود اوغير معروف الحجم بسبب كبره . فـأن :

$$\sigma_{\bar{p}} = \sqrt{\frac{P(1-P)}{n}} \qquad \ldots (16-4)$$

وعند عدم معرفة نسبة المجتمع (P) مع معرفة حجم المجتمع (N) فأن :

$$S_{\bar{p}} = \sqrt{\frac{N-n}{N-1}} \cdot \sqrt{\frac{P(1-P)}{n}} \qquad (17-4)$$

وفي حالة عدم معرفة نسبة المجتمع (P) وحجمه (N) فأن

$$S_{\bar{p}} = \sqrt{\frac{\bar{P}(1-P)}{n}} \qquad \ldots (18-4)$$

فاذا طبقنا المعادلة (4 – 15) على المثال 4 – 6 :

لحصلنا على مايلي :

$$\sigma_{p} = \sqrt{\frac{5-3}{5-1}} \cdot \sqrt{\frac{(0.4)(1.0-0.4)}{3}}$$

$$= 0.2$$

ويمكن للقاريء الحصول على نفس النتيجة فيما لو اتبع الأسلوب الأعتيادي لتحديد الأنحراف المعياري لنسب العينات العشرة التي أعتبرت مجتمعاً قائماً بذاته .

4-3-6 التوزيع العيني للمتوسط الحسابي للعينة ($\bar{X}$) :

ان التوزيع العيني للمتغير ($\bar{x}$) يتصف بمايلي :

(أ) : $u_{\bar{x}} = u_x$

(ب) : عند معرفة N فأن :

$$S_{\bar{x}} = \sqrt{\frac{N-n}{N-1}} \cdot \frac{S_x}{\sqrt{n}}$$

وعندما تكون N كبيرة جداً او غير معروفة بسبب الكبر فأن :

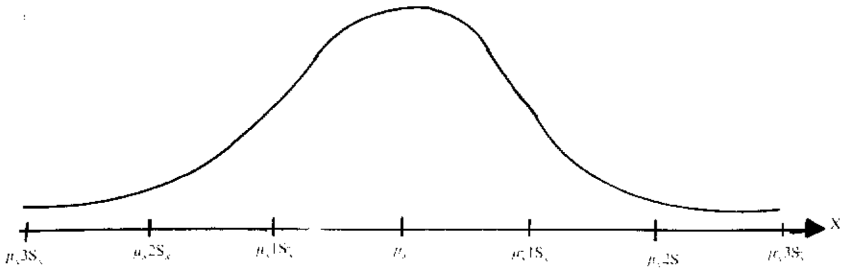
$$S_{\bar{x}} = \frac{S_x}{\sqrt{n}}$$

ج – اذا كان المجتمع المأخوذ منه العينة موزعاً توزيعاً معتدلاً ، فأن التوزيع العينـــي للمتغير $\bar{X}$ هو توزيع معتدل بغض النظر عن حجم العينة .

د – اذا كان حجم العينة كبيراً ، فأن التوزيع العيني للمتغير $\bar{x}$ هو توزيع معتدل تقريباً وهذه القاعدة تنطبق على جميع المجتمعات المأخوذة منها العينة .

ويلاحظ بأننا ركزنا على الحالة التي لاتتوفر فيها قيمة الأنحراف المعياري للمجتمع (σ_x) لأسباب ذكرناها سابقاً .

ويمكن توضيح هذا التوزيع في حالة كون حجم العينة كبيراً بما فيه الكفاية كما فـــي الشكل 4 – 8 .



شكل (4 - 8) : التوزيع العيني للمتغير $\bar{x}$ في حالة كون n كبيرا نسبيا

7-3-4 التوزيع العيني لنسبة العينة ($\bar{P}$)

يتصف التوزيع العيني للمتغير ($\bar{p}$) بمايلي :

: (أ) $\mu_{\bar{p}} = P$

(ب) $\sigma_{\bar{p}} = \sqrt{\frac{N-n}{N-1}} \cdot \sqrt{\frac{P(1-p)}{n}}$

او

$$S_{\bar{p}} = \sqrt{\frac{P(1-P)}{n}}$$

عندما يكون حجم المجتمع N كبيرا جدا وغير معروف .

او أن

$$S_{\bar{p}} = \sqrt{\frac{\bar{P}(1-\bar{P})}{n}}$$

في حالة عدم معرفة كل من P , N

ج – اذا كان المجتمع المأخوذة منه العينة موزعاً توزيعاً معتدلاً . فأن التوزيع العيني للمتغير P هو توزيع معتدل بغض النظر عن حجم العينة.

د – اذا كان حجم العينة كبيراً . فأن التوزيع العيني للمتغير ($\bar{P}$) هو توزيع معتدل تقريباً بغض النظر عن توزيع المجتمع الأصلي .

تماريـن

4 - 1 : اذا كانت نسبة الاصابة بمرض الكوكسيديا في أحد حقول الدواجن هـي 20 ٪ . فلوأخذت عينة عشوائية مكونة من 6 دجاجات ، فما هواحتمال

(أ) وجود دجاجتين مصابتين ؟
(ب) عدم وجود اي دجاجة مصابة ؟
(ج) وجود 4 دجاجات سليمة في الاقل ؟

4 - 2 : لو علمت ان وزن اللحم الصافي لدجاج اللحم المنتج من قبل مزرعة معينة موزع توزيعاً معتدلا بمتوسط حسابي 1·20 كغم وانحراف معياري 0·2 كغم . المطلوب :

(أ) رسم الشكل الذي يمثل هذا التوزيع .
(ب) ما هواحتمال كون وزن دجاجة مختارة بشكل عشوائي يساوي . اويقل عن 1·5 كغم ؟ وضح اجابتك بالرسم أيضاً .
(ج) ماهو أحتمال كون وزن دجاجة مختارة بشكل عشوائي يتراوح بين 0·8 و 1·5 كغم ؟
(د) لو أخذت عينة مكونة من 25 دجاجة ، فما هو احتمال كون المتوسط الحسابي لوزن الدجاجة في العينة يتراوح بين 1·0 كغم و 1·4 كغم ؟

4 - 3 : ارسم التوزيع المعتدل t مبيناً القيم التي تعزل مساحتين متساويتين من على جهتيه تمثلان 0·05 من المساحة الكلية اذا علمت ان درجات الحرية تساوي 19 .