المحور الثاني: نظرية التقدير

1-مقدمة في نظرية التقدير:

يمكن النظر إلى نظرية التقدير على أنها جزئين متكاملين، يهتم الجزء الأول بالبحث عن أفضل تقدير لمعلمة مجهولة تخص المجتمع، في حين يهتم الجزء الثاني بالبحث عن أفضل فترة يمكن حصر قيمة المعلمة المجهولة خلالها وغالبا ما يسمى "**التقدير بفترة**"، فعلى فرض أن ....X1, X2, X3, تمثل قياسات عينة عشوائية ذات حجم مسحوبة من مجتمع بدالة كتلة احتمالية $p(x, \theta)$ او دالة كثافة احتمالية (θ) حيث أن θ تمثل معلمة تخص وتشخص ذلك المجتمع، واننا نرغب وعلى أساس قياسات هذه العينة الوصول إلى أفضل تقدير ممكن إلى $\theta^{\wedge}$ مثل فمثلا إذا كانت هذه العينة مسحوبة من مجتمع في توزيع بواسون بالمعلمة λ فإن الأمر منصب على تقدير أفضل قيمة عددية ممكنة إلى $\lambda = \theta$ ، وإذا كانت هذه العينة مسحوبة من $N(\mu,\delta)$ فان الأمر سيكون منصب على تقدير أفضل قيمة عددية لكل من

$\theta_1=\mu$ و $\theta_2=\delta$، إن التقدير (أو المؤشر الإحصائي) θ عبارة عن دالة بدلالة قياسات العينة لا يعتمد على معلمة (أو معالم) التوزيع، أي أن$\theta^{\wedge} = g(x_1, x_2, x_3,........, x_n)$ وأن $\theta^{\wedge}$ يعتبر متغير عشوائي يسلك وفق دالة كتلة أو كثافة احتمالية معينة تعتمد على حجم العينة n والتوزيع الاحتمالي للمتغير العشوائي X.

- **خواص جودة التقدير:** تتلخص معايير دقة التقدير بما يلي:

- **عدم التحيز:** وهذا يشير إلى أن تقدير الوسط الحسابي $\bar{x}$ للعينة يعود إلى متغير عشوائي من حيث الحجم وعملية السحب من المجتمع مستوفيا لشروط العشوائية أي أن:

$E(\bar{x})=\mu$

حيث أن:

μ: يمثل متوسط المجتمع الذي سحبت منه العينة. فعند تساوي كلا الوسطين يكون التقدير غير متحيز؛

$\bar{x}$: متوسط العينة.

مثال: إن وسط العينة $\bar{x}$ تقدير غير متحيز لوسط المجتمع μ أما تباين العينة S^2 فهو تقدير متحيز لتباين المجتمعδ^2

البرهان:

$$E(\bar{x})= E(\frac{\sum X}{n})=\frac{1}{n}\sum E(x)= \frac{1}{n}\sum\mu=\frac{1}{n}(n\ \mu)= \mu$$

أي أن: $\bar{x}$ هو تقدير غير متحيز لـــ: μ

ومن أجل إثبات أن S^2 هو تقدير متحيز لتباين المجتمع δ^2 نكتب:

$$S^2= \frac{1}{n}\sum (x-\bar{x})^2= \frac{1}{n}\sum [(x-\mu) -(\bar{x}-\mu)]^2=\frac{1}{n}\sum [(x-\mu)^2 -n(\bar{x}-\mu)]^2$$

ولهذا فإن:

$$E(S^2)= \frac{1}{n} [\sum E(x-\mu)^2 - n\ E(\bar{x}-\mu)^2]= \frac{1}{n} [E\ \delta^2 - n\ \frac{\delta 2}{n}] =\frac{\delta 2}{n} (n-1) =\frac{n-1}{n}\ \delta^2\neq \delta^2$$

أي أن تباين العينة S^2 ليس تقديراً غير متحيز لتباين المجتمع δ^2.

- **التناسق:** وهو قرب قيمة التقدير إلى قيمة المعلمة، فكلما ازداد حجم العينة يكون التقدير قد حقق التناسق مع العينة؛
- **الكفاءة:** وتشير إلى التغاير في العينة المستخدمة لأي تقدير، إن التقدير الأقل تباين يكون نسبياً الأكثر كفاءة.

$$\theta\hat{}(x) = \frac{n}{\sum \mathrm{xi}}$$

2-طرق حساب التقدير: غالبا ما تكون المجتمعات التي نريد معرفة خصائصها مجهولة المعالم كالوسط الحسابي μ، والانحراف المعياري δ، وهناك من الأسباب العلمية والاقتصادية ما يحول دون تحديد هذه الخصائص تحديدا مؤكدا ودقيقا، وفي هذه الحالة فإننا نلجأ إلى سحب عينة عشوائية من هذا المجتمع ثم نحسب إحصائيات هذه العينة مثل: $\overline{X}, S, p$ محاولين تقدير معالم المجتمع الأصلي: μ,δ,p من خلال إحصائيات هذه العينة. وهذا ما سنقوم بتفصيله الآن حيث نقسم طرق التقدير إلى نوعين: **التقدير بنقطة، والتقدير بفترة (التقدير بمجال ثقة).**

أولا: التقدير بنقطة

بعد أن تعرفنا إلى **خواص المقدرات** نستعرض أهم طريقتين مستخدمتين في إيجاد المقدرات، وهما: طريقة الإمكانية العظمى تسمى أحيانا طريقة الإمكان الأكبر، وطريقة العزوم، وتعتبر طريقة الإمكانية العظمى من أهم طرق التقدير النقطي وخاصة في حالة العينات الكبيرة وتعزى هذه الطريقة لرونالد فيشر (R.Fisher)، وهي تعطي مقدرات كافية إن وجدت إلا أنها قد تكون متحيزة أحيانا، ولكن في الحالة التي يؤول فيها حجم العينة n إلى ما لانهاية، فإنها تعطي مقدرات غير متحيزة، ولها أقل تباين، ومنسقة دوماً، كما يؤول توزيعها إلى التوزيع الطبيعي مع زيادة حجم العينة.

بفرض أن دالة كثافة الاحتمال $f(x,\theta)$ حيث θ تمثل المعلمة المراد تقديرها، وبفرض أن متغيرات القيم الملاحظة في العينة من الشكل : $x_1,x_2,x_3,.....,x_n$، عندئذ يكون تابع الكثافة المشترك للعينة ولنرمز له بالرمز: $L(x,\theta)$ بالشكل التالي:

$$\pi\, f(x, \theta)= f(x_1, \theta).\, f(x_2, \theta).\, f(x_3, \theta)\ldots\ldots\ldots\, f(x_n, \theta)$$

ويعني التقدير بنقطة أن نعبر عن معلمة المجتمع المجهولة بقيمة واحدة يشارك كل أو بعض مفردات أو قيم العينة في تكوينها، بحيث إذا تحققت الشروط الخاصة بالتقدير الجيد

يمكن اعتبار هذه القيمة تقديرا جيدا لمعلمة المجتمع، وما يعاب على هذه الطريقة أنه في كثير من الأحيان تكون قيمة التقدير غير مطابقة لقيمة المعلمة المجهولة للمجتمع، الأمر الذي لا نستطيع معه معرفة مدى دقة التقدير أو مدى بعده عن القيمة الحقيقية المجهولة المعلمة المجتمع. هذا ويمكن اعتبار أن:

- متوسط العينة $\bar{X}$ تقدير نقطة جيد المتوسط المجتمع μ.
- الفرق $\bar{X}-\bar{Y}$ تقدير نقطة جيد للفرق $\mu_1-\mu_2$
- النسبة p^ تقدير نقطة جيد للنسبة p
- الفرق p_1-p_2 تقدير نقطة جيد للفرق p_1-p_2

وهذا يعني إعطاء قيمة محددة لمعلمة المجتمع من خلال إحصائيات العينة، وعند إتباع هذا الأسلوب الإحصائي في التقدير فإنه يجب مراعاة شروط التقدير الجيد ومحاولة توفر أغلبية هذه الشروط أو الخصائص حتى يمكن التوصل إلى أدق التقديرات لمعالم المجتمعات إحصائيا.

فإذا كان متوسط إنتاجية العمال في عينة من 100 عامل هو 170 وحدة يوميا فإنه يمكن اعتبار هذا المتوسط تقدير غير متحيز ومتسق لمتوسط المجتمع، ونستطيع القول في هذه الحالة أن متوسط المجتمع 170 وحدة يوميا .

وهناك طرق أخرى لتقدير معالم المجتمع مثل: **طريقة المربعات الصغرى، طريقة العزوم، وطريقة الإمكان الأعظم**، وأخيرا **طريقة الرسم**.

- **طريقة المعقولة العظمى:**

تدعى القيمة (النقطة) (x) θ بتقدير المعقولية العظمى للمعلمة إذا كانت دالة المعقولية (θ) تبلغ نهايتها العظمى عندها أي أن:

$L(x; \theta\hat{}) \geq L(x; \theta); \quad \forall\, \theta \in \theta$

أو:

$$L(x; \theta\hat{}) = \sup_{\theta \in \theta} L(x; \theta)$$

وهذا يعني أن التقدير ^θ في حالة التوزيعات المنقطعة عبارة عن قيمة و التي تجعل احتمال سحب العينة المشاهدة x (قيمة لــ X) أكبر ما يمكن، أما بالنسبة للتوزيعات المستمرة فإن قيمة θ والتي تجعل احتمال الحصول على مفردات العينة X قريبة جداً من القيم التي حصلنا عليها (القيم $x_1, x_2, x_3, \ldots\ldots$) أكبر ما يمكن

يحقق تقدير المعقولة العظمى $\theta\hat{}=\theta\hat{}(x)$ المعادلات التالية:

1- $L(X, \theta)= \pi\, f(x, \theta)= f(x_1, \theta).\, f(x_2, \theta).\, f(x_3, \theta)\ldots\ldots\ldots\ldots f(x_n, \theta)$
2- $\text{Ln}\, L(X, \theta)$
3- $\frac{\partial L(X,\theta)}{\partial \theta}$
4- $\frac{\partial L(X,\theta)}{\partial \theta} = 0$

مثال: إذا كانت $X = (x_1, x_2, x_3, \ldots\ldots\ldots)$ عينة عشوائية من مجتمع كثافة توزيعه الاحتمالية كالتالي:

$f(x,\theta)= \theta\, e^{-\theta xi} \qquad ; x > 0 \ , \theta > 0$

المطلوب: أوجد مقدر المعقولية العظمي للمعلمة θ ؟

الحل:

1- $L(X, \theta)= \pi\ f(x, \theta)= \theta^{n}\ e^{-\theta\sum xi}$
2- $Ln\ L(X, \theta)= n \ln \theta - \theta \sum xi$
3- $\frac{\partial L(X,\theta)}{\partial \theta}= \frac{n}{\theta} - \sum xi$
4- $\frac{\partial L(X,\theta)}{\partial \theta}= \frac{n}{\theta} - \sum xi = 0$

وعليه فإن: $\theta\hat{} = \frac{1}{\bar{x}}$

ثانيا: التقدير بفترة ثقة

هناك العديد من الأسباب التي دعت الإحصائيين إلى تقدير فترة تقع داخلها معلمة المجتمع المراد الحصول عليها، فالتقدير بنقطة لا يقترن بدرجة معينة من الصحة أو الثقة في النتائج، بالإضافة إلى صعوبة تحقيق الخصائص الواجب توافرها في المقدر الجيد، ناهيك عن أن أخذ العينات ينتج عنه أخطاء من أنواع مختلفة لذلك كله فمن الأفضل ربط تقدير المعلمة المجهولة بدرجة معينة للثقة في النتائج أو بنسبة معينة للخطأ في التقدير ووضع حدين، حد أعلى (يسمى بالحد الأعلى للثقة وحد أدنى يسمى (بالحد الأدنى للثقة)، والفرق بينهما يسمى بمدى الثقة. وفي إطار هذا المدى تقع معلمة المجتمع عند مستوى الثقة المحددة سلفا.

قاعدة عامة: إذا كان لدينا مجتمع يتكون من N وحدة معلمته المجهولة θ، وإذا سحبنا منه عينة عشوائية حجمها n وحسبنا قيمة θ منها، فإنه باستخدام نظرية النهاية المركزية نجد أن المتغير:

$$Z=\frac{\theta\hat{}-E(\theta\hat{})}{\sqrt{var}(\theta\hat{})}$$

فالتقدير النقطي للوسط (μ) ينحصر بحساب قيمة أعلى خطأ (ε) في التقدير. وكثيراً ما تختلف قيمة الوسيط عن ذلك وعلى هذا الأساس يفضل في كثير من الأحيان استبدال التقدير النقطي للوسط (μ) بالتقدير المجالي أو كما يصطلح عليه بفترة الثقة حيث يمكننا التأكيد

وبدرجة معقولة من الثقة أو بتعبير آخر بدرجة احتمال عالية P بأن الوسط (μ) يقع ضمن فترة التقديرات المحسوبة وقد لاحظنا في هذا الصدد بأن مقدار الخطأ النسبي يقع ضمن الفترة التالية:

$$-Z_{\alpha/2} < \frac{\bar{X}-\mu}{\frac{\delta}{\sqrt{n}}} < +Z_{\alpha/2} \quad \text{...(1)}$$

وذلك عندما يكون انحراف المجتمع (δ) معلوما حيث درجة الثقة تكون: $p = (1-\alpha)\ \%$

وبترتيب العلاقة أعلاه نجد أن مجال الثقة لمتوسط المجتمع (μ) يأخذ الشكل التالي:

$$\bar{X} - Z_{\alpha/2}\,\delta_{\bar{x}} < \mu < \bar{X} + Z_{\alpha/2}\,\delta_{\bar{x}} \quad \text{...(2)}$$

ونطلق على المتراجحة الأخيرة اسم **مجال الثقة الطبيعي لتقدير متوسط المجتمع الإحصائي (μ)**

ملاحظة:

وتبقى العلاقة (2) المستخدمة في الحالة التي يكون فيها انحراف المجتمع (δ) مجهول وحجم العينة كبيرا ($n \geq 30$) حيث نستبدل انحراف المجتمع الإحصائي (δ) بانحراف العينة العشوائية (S) أما في الحالة التي يكون فيها (δ) مجهول وحجم العينة ($n<30$) والمجتمع يتبع التوزيع الطبيعي فإننا نجد من نظرية التوزيعات التكرارية للعينات بأن الخطأ النسبي ($^{\varepsilon}$) يقع ضمن الفترة التالية:

$$-t_{\alpha/2} < \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} < +t_{\alpha/2} \quad \text{...(3)}$$

وبدرجة ثقة $p(1-\alpha)\ \%$ ودرجة حرية $r = n-1$ ومن العلاقة (3) نجد أن التقدير المجالي للوسط (μ) يأخذ الشكل التالي:

$$\bar{X} - t_{\alpha/2}\frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2}\frac{S}{\sqrt{n}} \quad \text{...(4)}$$

ونطلق على المتراجحة الأخيرة اسم مجال أو فترة Student (t) لتقدير (μ)

يخضع للتوزيع الطبيعي باحتمال: $1-\alpha$ وعليه فإن:

$$P\left(\hat{\theta} - Z_{\alpha/2}\sqrt{var(\hat{\theta})} < \theta < \hat{\theta} + Z_{\alpha/2}\sqrt{var(\hat{\theta})}\right) = 1-\alpha$$

حيث أن:

- $(\theta\hat{}) \sqrt{var}\, \mathbf{Z}_{\alpha/2}$ -^ **θ:** الحد الأدنى للثقة
- $(\theta\hat{}) \sqrt{var}\, \mathbf{Z}_{\alpha/2}$ +^ **θ:** الحد الأعلى للثقة
- $\mathbf{Z}_{\alpha/2}$: القيمة المعيارية المستخرجة من جداول المنحنى المعتدل المناظرة لاحتمال $\alpha/2$ بمعنى المساحة المحصورة ما بين $Z_{\alpha/2}$ ونهاية التوزيع المعتدل $\alpha/2$
- $1-\alpha$: مستوى الثقة
- α: نسبة الخطأ في التقدير
- $\sqrt{var\ (\theta\hat{})}$: الانحراف المعياري لـــ: $\theta\hat{}$

وبشكل عام تكون فترة الثقة على الصيغة التالية:

تقدير نقطة $\pm$ (معامل الثقة)(الخطأ المعياري)

مثال 01:

أستخدمت عينة عشوائية من مائة محرك لدراسة قدرة مجتمع إحصائي من المحركات ذي انحراف معياري $\delta = 5.1$ حصان، فإذا كان المتوسط الحسابي للعينة المأخوذة من المجتمع هو $\bar{X} = 21.6$ حصان

المطلوب: أوجد فترة الثقة لقيم μ المحتملة وبدرجة ثقة 95 %؟

الحل:

$n\geq100$، $\delta = 5.1$ (معلوم)، بما أن: $(1-\alpha)=0.95$ فإن: $Z_{\alpha/2} = 1.96$

$$p\,(\bar{X}- Z_{\alpha/2}\,\delta_{\bar{x}} < \mu < \bar{X}+ Z_{\alpha/2}\,\delta_{\bar{x}}) = 1-\alpha$$

$$p(21.6- 1.96\,\frac{5.1}{\sqrt{100}} < \mu < 21.6+ 1.96\,\frac{5.1}{\sqrt{100}}) = 0.95$$

$$p(20.6 < \mu < 22.6) = 0.95$$

ويعني ذلك بأننا على ثقة بدرجة 0.95 بأن الفترة (المجال) أعلاه تحتوي على قيمة (μ) أو أننا على ثقة بدرجة 95 % بأن: (μ) لا تتعدى الحد الأعلى: $\mu = 22.6$ ولا تقل عن الحد الأدنى: $\mu = 20.6$

مثال 02:

إذا كان متوسط الوزن المفقود لعينة من أحجار الفرامل متكونة من 16 حجرا بعد مسافة 200 كلم هو $\overline{X} = 3.42$، وكان الانحراف المعياري لعناصر العينة يساوي 0.68.

المطلوب: ما هو تقدير المتوسط الحسابي للوزن المفقود للمجتمع الإحصائي (أحجار الفرامل) وبدرجة ثقة 99 %؟

الحل:

$n= 16< 30$، $\delta = ?$ (مجهول)، $\alpha =0.01$

بما أن توزيع الوزن المفقود طبيعي فإننا نستخدم توزيع Student (t) لتحديد تقدير متوسط المجتمع μ

- بمقاطعة 0.005 أفقيا مع درجة الحرية v=15 في جدول Student نجد:

$t_{0.005} = 2.947$

$$p\left(\overline{X}- t_{\alpha/2}\frac{S}{\sqrt{n}}< \mu <\overline{X}+ t_{\alpha/2}\frac{S}{\sqrt{n}}\right)=1- \alpha$$

بالتطبيق العددي نجد:

$$p\left(3.42- (2.947)\frac{0.68}{\sqrt{16}}< \mu <3.42+ (2.947)\frac{0.68}{\sqrt{16}}\right)=0.99$$

$$p\left(3.42- (2.947)\frac{0.68}{\sqrt{16}}< \mu <3.42+ (2.947)\frac{0.68}{\sqrt{16}}\right)=0.99$$

$$p(2.92< \mu <3.92)=0.99$$

ويعني ذلك وجوب تغيير أحجار الفرامل بعد مسافة 200 كلم بحد أدنى للوزن المفقود منها يساوي: 2.92 غ، وحد أعلى للوزن المفقود منها يساوي: 3.92 غ.

3–تقدير حجم العينة:

في كثير من الأحيان يراد سحب عينة عشوائية من مجتمع لتقدير الوسط الحسابي للمجتمع بخطأ لا يزيد عن عدد معين. في هذه الحالة لابد من تعيين عدد عناصر العينة مسبقا.

يمتد المجال في العلاقة رقم (2) بين النقطتين (الحدين):

$\overline{X}+ Z_{\alpha/2}\, \delta_{\overline{x}}$ ،$\overline{X}- Z_{\alpha/2}\, \delta_{\overline{x}}$

وتقع $\overline{X}$ في وسطه، ولذلك إذا اعتبرنا أن: $\overline{X}$ تقدير نقطي لـــ μ فإن الخطأ المرتكب هو:

$Z_{\alpha/2}\, \delta_{\overline{x}}$بمستوى ثقة: $(1-\alpha)$

لنرمز بـــ: E للحد الأعلى للخطأ المسموح به في تقدير μ (أو دقة تقدير متوسط المجتمع μ)، ولتعيين عدد عناصر العينة (n) يجب أن لا يتجاوز الحد الأعلى للخطأ المسموح به: E بمستوى ثقة: $(1-\alpha)$ ونكتب:

$P\,(Z_{\alpha/2}\, \delta_{\overline{x}} \leq E)=1-\alpha$

بتربيع الطرفين نجد:

$Z^2_{\alpha/2}\ \frac{\delta 2}{n} \leq E$

ومنه: $n \geq \left(\frac{Z_{\frac{\alpha}{2}}\delta}{E}\right)^2$

مثال 01:

نريد تقدير الوسط الحسابي المجتمع بخطأ لا يزيد عن 5 بمستوى ثقة 90 %، عن طريق حساب الوسط الحسابي لعينة عشوائية من المجتمع.

المطلوب:

كم ينبغي أن يكون عدد عناصر هذه العينة إذا علمت أن الانحراف المعياري للمجتمع يساوي 20.

الحل:

$$Z_{\alpha/2} = Z_{0.05} = 1.645$$

$$n \geq \left(\frac{Z_{\frac{\alpha}{2}}\delta}{E}\right)^2$$

$$n \geq \left(\frac{1.645(20)}{5}\right)^2 \approx 43.29$$

وبالتالي فإن أقل حجم يمكن للعينة يجب أن يكون: n= 44

مثال 02:

في مجتمع يتكون من 500 وحدة، أوجد حجم العينة اللازم لتقدير الوسط الحسابي وذلك باحتمال 95 % وبدقة تساوي 0.9، إذا كان الانحراف المعياري للمجتمع يساوي 7 .

الحل:

$$Z_{\alpha/2} = Z_{0.025} = 1.96$$

$$n \geq \left(\frac{Z_{\frac{\alpha}{2}}\delta}{E}\right)^2 \geq \left(\frac{1.96\,(7)}{0.9}\right)^2 \geq 232$$

وبالتالي فإن أقل حجم يمكن للعينة يجب أن يكون: n= 232