The Maximum Likelihood (ML) Estimator

Ex The probability $p = P(H)$ of a coin may be 0.1 or it may be 0.9. To resolve the uncertainty, the coin was tossed 10 times and 3 heads were observed. Find a maximum likelihood estimate for p.

$\to$ probability of success

$p = 0.1$ or $p = 0.9$ binomial أنو لاحظ +

X: number of heads observed is a binomial R.V.

$$P(X=x) = \binom{n}{x} p^x [1-p]^{n-x} \quad , n=10$$

$$\therefore P(X=x) = \binom{10}{x} p^x [1-p]^{10-x}$$

$$P(X=3 \overset{as}{;} p=0.1) = \binom{10}{3} (0.1)^3 (0.9)^7 = 0.0574$$

But, when:-

$$P(X=3; p=0.9) = \binom{10}{3} (0.9)^3 (0.1)^7 = 8.748 \times 10^{-6}.$$

AS whe $\boxed{p = 0.1}$ has the higher probability, thus we choose it as our estimator. $\therefore \boxed{\hat{p} = 0.1}$

B if $\boxed{X = 8}$; $P(X=8; p=0.1) = \binom{10}{8} (0.1)^8 (0.9)^2 = 3.645 \times 10^{-}$

$$P(X=8; p=0.9) = \binom{10}{8} (0.9)^8 (0.1)^2 = 0.1937.$$

In this case, we choose $\boxed{\hat{p} = 0.9}$ as the estimator.

Example Let p be the probability of success in a binomial distribution. This probability is unknown. To estimate p, the experiment is performed 10 times and 3 successes were observed. Find a maximum likelihood estimate for p.

binomial

$n = 10$, $X = 3$

$p = ??$ → we find it

$f(p) = P(X=3\,;\,p) = \binom{10}{3} p^3 (1-p)^7 = \frac{10!}{3!\,(10-7)!}\, p^3 (1-p)^7$

بالضبط P ونساويها بالصفر.

عشان نوجد أعلى قيمة للـ success ، بنشتق المعادلة

$\therefore \frac{df}{dp} = \binom{10}{3}\left[-7p^3(1-p)^6 + 3p^2(1-p)^7\right] = 0$

وبيطلع معي إشي 0.25 , 0.35 , و $P = 0.3$

ونعوض بهاي المعادلة وبشوف أي نتيجة كانت فيها البروبابلتي أعلى

والقيمة هاي هي بتكون $\hat{p}$ ، فهذا اللي طلع معنا

$\boxed{\hat{p} = \frac{3}{10} = 0.3}$

For the continuous case :- بنكتب ال cdf الخاص بالفنكشن
زي ما كتبنا ال pdf بال Discrete .

Example (6-3) Given a random sample of size (n) taken from a Gaussian population with parameters μ_x and σ_x^2. Use the ML technique to find estimators for the cases :-

a The mean μ_x when the variance σ_x^2 is assumed known

Gaussian μ_x , σ_x^2 → random sample size n → $x_1, x_2, \dots, x_n$

$$f_{x_1}(x_1) = \frac{1}{\sqrt{2\pi\sigma_{x_1}^2}} e^{\frac{-(x_1-\mu_{x_1})^2}{2\sigma_{x_1}^2}}$$

بما أن $\mu_{x_1} = \mu_x$ و $\sigma_{x_1}^2 = \sigma_x^2$ ∴

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x_1-\mu_x)^2}{2\sigma_x^2}}$$

بس $x_1 \neq x$ يعني طبعا إذا كان $x_1 = z$ ما بمشي ال x بال z .

And $$f_{x_2}(x_2) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{\frac{-(x_2-\mu_x)^2}{2\sigma_x^2}}$$

$$\vdots$$

$$\therefore f_{x_n}(x_n) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{\frac{-(x_n-\mu_x)^2}{2\sigma_x^2}}$$

لما أنه إحنا ماخذين random sample فالعينة independent من بعض وكل واحد لا يعتمد على الأخرى

$$\therefore f_{x_1,x_2,\dots,x_n}(x_1,x_2,\dots,x_n) = f_{x_1}(x_1) f_{x_2}(x_2) \dots f_{x_n}(x_n)$$

$$= \left[\frac{1}{\sqrt{2\pi\sigma_x^2}}\right]^n e^{\frac{-1}{2\sigma_x^2}\sum_{i=1}^{n}(x_i-\mu_x)^2}$$

حطينا مجموع علشان الضرب ع جمع القوى .

$$\ln\left[f_{x_1,\dots,x_n}(x_1,\dots,x_n)\right] = \ln\left[(2\pi\sigma_x^2)^{\frac{-n}{2}} e^{\frac{-1}{2\sigma_x^2}\sum_{i=1}^{n}(x_i-\mu_x)^2}\right]$$

$$L(\mu_x, \sigma_x^2) = \frac{-n}{2}\ln(2\pi\sigma_x^2) - \frac{1}{2\sigma_x^2}\sum_{i=1}^{n}(x_i-\mu_x)^2$$

صار اسمه
ML function
Maximum Likelihood

هلأ بلّش حل $\boxed{a}$ حيث إنو رح أستخدم المعادلة السابقة في إيجاد الـ $\hat{\mu}_x$ المقدّر، على فرض إني بعرف σ_x^2.

→ $\boxed{a}$ Find $\hat{\mu}_x$, σ_x^2 is known (provided that):-

$$L(\hat{\mu}_x) = \frac{-n}{2}\ln(2\pi\sigma_x^2) - \frac{1}{2\sigma_x^2}\sum_{i=1}^{n}(x_i-\hat{\mu}_x)^2$$

هون بنحط الإشي اللي بدي أوجد له estimation، وبنشتق المعادلة بالنسبة له.

$$\frac{dL(\hat{\mu}_x)}{d\hat{\mu}_x} = \underbrace{0}_{\text{مشتقتها بالنسبة إلى } \hat{\mu}_x} - \frac{1}{2\sigma_x^2}\sum_{i=1}^{n}2(x_i-\hat{\mu}_x)(-1) = 0$$

$$\therefore \frac{-1}{2\sigma_x^2}\sum_{i=1}^{n}2(x_i-\hat{\mu}_x)(-1) = 0$$

$$\sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\hat{\mu}_x = 0$$

$$\therefore \sum_{i=1}^{n}x_i = n\hat{\mu}_x \Rightarrow \boxed{\hat{\mu}_x = \frac{\sum x_i}{n}}$$

← ML estimator for the mean when the variance is known.

Now we check for biasing:-

$$E\{\hat{\mu}_x\} = E\left\{\frac{1}{n}\sum_n x_i\right\} = \frac{1}{n}\sum_{i=1}^{n}E\{x_i\} = \frac{1}{n}\sum_{i=1}^{n}\mu_x$$

$$= \frac{1}{n}(n)\mu_x = \mu_x$$

$\boxed{E\{\hat{\mu}_x\} = \mu_x}$ so it is unbiased.

$\boxed{a_2}$ $\hat{\mu}_x$, Variance is unknown.

$$L(\hat{\mu}_x, \hat{\sigma}_x^2) = \frac{-n}{2}\ln(2\pi\hat{\sigma}_x^2) - \frac{1}{2\hat{\sigma}_x^2}\sum_{i=1}^{n}(x_i-\hat{\mu}_x)^2$$

$$\frac{dL(\hat{\mu}_x, \hat{\sigma}_x^2)}{d\hat{\mu}_x} = 0 + \cancel{\frac{1}{2\hat{\sigma}_x^2}}\sum_{i=1}^{n}\cancel{2}(x_i-\hat{\mu}_x)\cancel{(-1)} = 0$$

$$\therefore \sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\hat{\mu}_x = 0 \Rightarrow \boxed{\hat{\mu}_x = \frac{1}{n}\sum_{i=1}^{n}x_i}$$

← ML estimator for the mean, when the variance is unknown.

يتبع

Check ~~for~~ biasing: $E\{\hat{\mu}_x\} = \ldots = \mu_x$

So it is biased

زي الفرع اللي قبله يعني .

$\boxed{b}$ Estimator for σ_x^2, μ_x is known.

$$L(\hat{\sigma}_x^2) = \frac{-n}{2}\ln(2\pi\hat{\sigma}_x^2) - \frac{1}{2\hat{\sigma}_x^2}\sum_{i=1}^{n}(x_i-\mu_x)^2$$

$$\frac{dL(\hat{\sigma}_x^2)}{d\hat{\sigma}_x^2} = \left(-\frac{n}{2} * \frac{1}{2\pi\hat{\sigma}_x^2} * 2\pi\right) + \frac{1}{2}\left(\frac{+1}{(\hat{\sigma}_x^2)^2}\right)\sum_{i=1}^{n}(x_i-\mu_x)^2 = 0$$

$$\frac{+n}{\cancel{2}\hat{\sigma}_x^{\cancel{2}}} = \frac{1}{\cancel{2}(\hat{\sigma}_x^2)^{\cancel{2}}}\sum_{i=1}^{n}(x_i-\mu_x)^2$$

$$\hat{\sigma}_x^2 = \frac{\sum(x_i-\mu_x)^2}{n}$$ → ML estimator for the variance when the mean is known.

→ check for biasing:-

$$E\{\hat{\sigma}_x^2\} = E\left\{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_x)^2\right\}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\underbrace{E\{(x_i-\mu_x)^2\}}$$

$$= \frac{1}{\cancel{n}}\sum_{i=1}^{n}\sigma_x^2 = \frac{1}{\cancel{n}} * \cancel{n}\,\sigma_x^2 = \sigma_x^2$$

So it is an unbiased estimator for the variance when the mean is known.

$\boxed{c}$ $\hat{\sigma}_x^2$ when μ_x is unknown.

$$L(\hat{\mu}_x, \hat{\sigma}_x^2) = \frac{-n}{2}\ln(2\pi\hat{\sigma}_x^2) - \frac{1}{2\hat{\sigma}_x^2}\sum_{i=1}^{n}(x_i-\hat{\mu}_x)^2$$

هون بنحط أكثر الاشي unknown

$$\frac{dL(\hat{\mu}_x, \hat{\sigma}_x^2)}{d\sigma_x^2} = \frac{-n}{2} * \frac{1}{\cancel{2\pi}\hat{\sigma}_x^2} * 2\pi - \frac{1}{2} * \frac{-1}{(\hat{\sigma}_x^2)^2}\sum_{i=1}^{n}(x_i-\hat{\mu}_x)^2 = 0$$

هون بنشتق حسب بدون estimator لنشتق بالسؤال

$$\frac{n}{2} * \frac{1}{\hat{\sigma}_x^2} = \frac{1}{2} * \frac{1}{(\hat{\sigma}_x^2)^2} \sum_{i=1}^{n} (x_i - \hat{\mu}_x)^2$$

$$\therefore \boxed{\hat{\sigma}_x^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu}_x)^2} \rightarrow$$ ML estimator for the variance when the mean is unknown.

→ Check for biasing:-

$$E\{\hat{\sigma}_x^2\} = E\left\{\frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu}_x)^2\right\}$$

$$= E\left\{\frac{1}{n} \sum_{i=1}^{n} (x_i^2 - 2x_i\hat{\mu}_x + \hat{\mu}_x^2)\right\}$$

$$= \frac{1}{n} E\left\{\sum x_i^2 - 2\hat{\mu}_x \underbrace{\sum_{i=1}^{n} x_i}_{n\hat{\mu}_x} + \sum \hat{\mu}_x^2\right\}$$

$$= \frac{1}{n} E\left\{\sum x_i^2 - \cancel{2}\hat{\mu}_x n\hat{\mu}_x + \cancel{n\hat{\mu}_x^2}\right\}$$

$$= \frac{1}{n} E\left\{\sum x_i^2 - n\hat{\mu}_x^2\right\}$$

$$= \frac{1}{n}\left[\sum E\{x_i^2\} - nE\{\hat{\mu}_x^2\}\right]$$

$$= \frac{1}{n}\left[\sum_{i=1}^{n} [\sigma_x^2 + \mu_x^2] - n\left[\frac{\sigma_x^2}{n} + \mu_x^2\right]\right]$$

$$* \ \sigma_x^2 = E\{x^2\} - \mu_x^2$$

$$\therefore E\{x^2\} = \sigma_x^2 + \mu_x^2$$

$$* \ Var\{\hat{\mu}_x\} = \frac{\sigma_x^2}{n}$$

$$\therefore E\{\hat{\mu}_x^2\} - (E\{\hat{\mu}_x\})^2 = \frac{\sigma_x^2}{n}$$

$$\therefore E\{\hat{\mu}_x^2\} = \frac{\sigma_x^2}{n} + \mu_x^2$$

$$= \frac{1}{n}\left[n\sigma_x^2 + \cancel{n\mu_x^2} - \sigma_x^2 - \cancel{n\mu_x^2}\right]$$

$$= \frac{n-1}{n}\sigma_x^2.$$

* Check for biasing:-

$$E\{\hat{\sigma}_x^2\} = E\left\{\frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu}_x)^2\right\} = \frac{n-1}{n}\sigma_x^2$$

علشان نخليه unbiased هنضربه ب $\frac{n}{n-1}$ كدا هنخلي σ_x^2 لوحدها ويصير unbiased

$$\therefore E\{\hat{\sigma}_{x_{new}}^2\} = E\left\{\frac{n}{n-1} * \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu}_x)^2\right\} = \frac{\cancel{n}}{n-1} * \frac{n-1}{\cancel{n}}\sigma_x^2$$

$$\therefore \boxed{\hat{\sigma}_{x_{new}}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu}_x)^2} \rightarrow$$ An unbiased ML estimator of the variance, when the mean is unknown.

Ex 6-4 Given a random sample of size (n) taken from a distribution X with pdf :-

$$f(x) = (\alpha+1)\, x^{\alpha} \quad , 0 < x < 1 .$$

Use ML technique to find an estimator for α.

The likelihood function is:-

$$L(\alpha) = f(x_1)\, f(x_2)\, f(x_3) \dots f(x_n)$$

وأخذناها كعامل مشترك

$$L(\alpha) = (\alpha+1)\, x_1^{\alpha} \dots (\alpha+1)\, x_n^{\alpha} = (\alpha+1)^n\, x_1^{\alpha} \dots x_n^{\alpha}$$

Enter Ln to both sides:-

$$\ln[L(\alpha)] = n \ln(\alpha+1) + \alpha \ln x_1 + \alpha \ln x_2 + \dots \alpha \ln x_n$$

$$\ln[L(\alpha)] = n \ln(\alpha+1) + \alpha \sum_{i=1}^{n} \ln(x_i)$$

$$\therefore \frac{d \ln[L(\alpha)]}{d\alpha} = n \frac{1}{1+\alpha} + \sum_{i=1}^{n} \ln(x_i) = 0$$

المعكوس

$$\frac{-n}{1+\hat{\alpha}} = \sum_{i=1}^{n} \ln(x_i) \Rightarrow \frac{1+\hat{\alpha}}{-n} = \frac{1}{\sum_{i=1}^{n} \ln(x_i)}$$

$$1+\hat{\alpha} = \frac{-n}{\sum_{i=1}^{n} \ln(x_i)} \Rightarrow \boxed{\hat{\alpha} = \frac{-n}{\sum_{i=1}^{n} \ln(x_i)} - 1}$$

هادا ال estimator

المطلوب بالسؤال

Finding Interval Estimators for the mean and Variance

$$P(\theta_1 \leq \theta \leq \theta_2) = 1-\alpha \; ; \; 0 < \alpha < 1$$

Where:- (θ) is the unknown parameter.

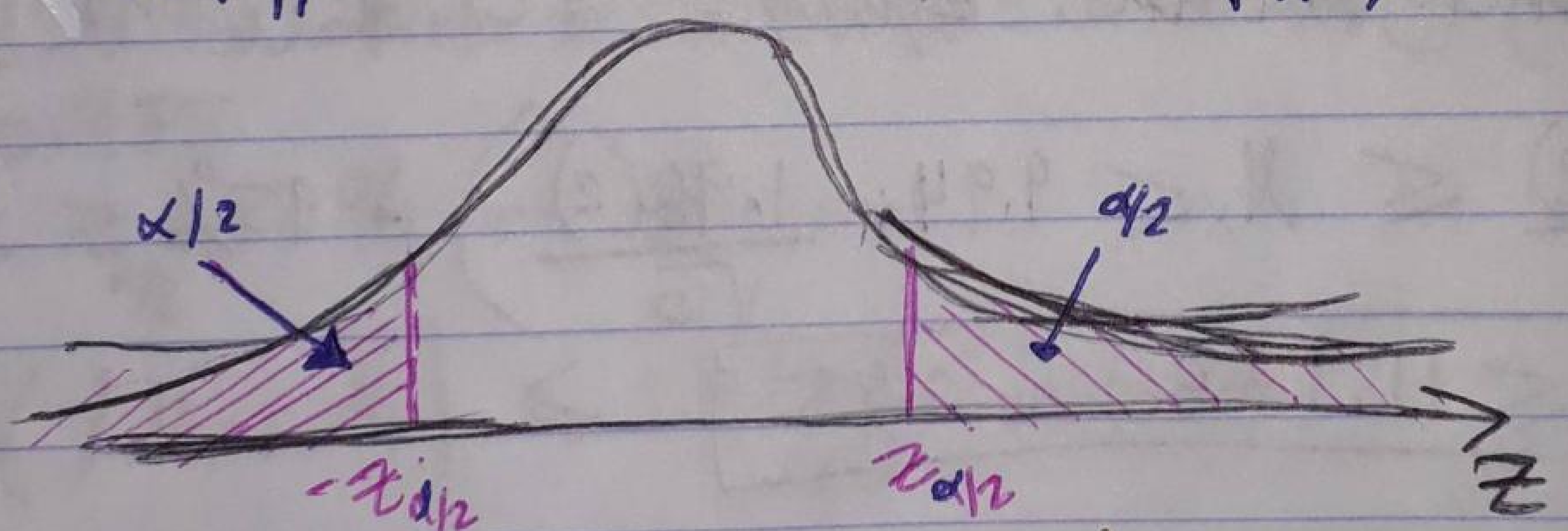
$(1-\alpha)$ is the confidence coefficient

(α) is the confidence level.

(θ_1) and (θ_2) are the lower and the upper confidence limits.

$\boxed{I}$ Confidence Interval on the Mean: (Known Variance).

$$P\left(\hat{\mu}_x - Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n}} \leq \mu_x \leq \hat{\mu}_x + Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n}}\right) = 1-\alpha$$



$\boxed{Ex}$ The following samples are drawn from a population that is known to be Gaussian.

7.31	10.80	11.27	11.91	5.51	8.00	9.03	14.42	10.24	10.91

Find the confidence limits for 95% confidence level if the variance of population is 4.

Gaussian ✓
Variance known ✓

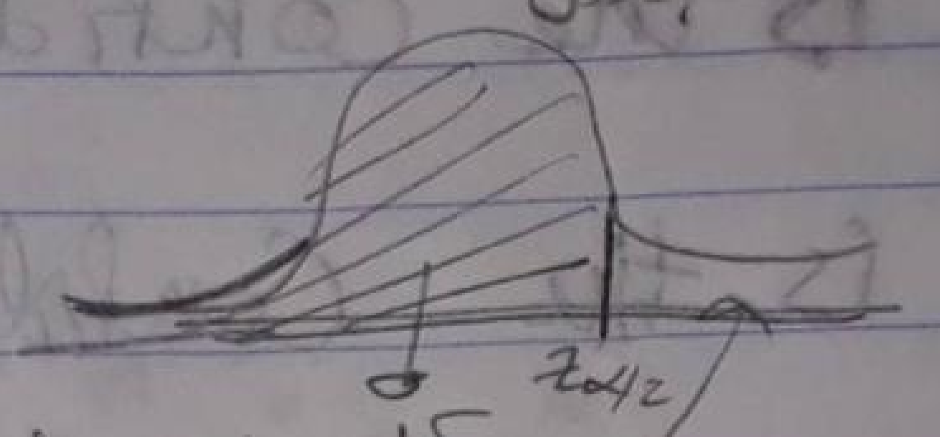
$\therefore$

$$P\left(\hat{\mu}_x - Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n}} \le \mu_x \le \hat{\mu}_x + Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n}}\right) = 1-\alpha$$

$\boxed{n = 10}$ $\boxed{1-\alpha = 0.95}$ $\boxed{\sigma_x^2 = 4}$ $\boxed{\hat{\mu}_x = \frac{1}{n}\sum_{i=1}^{n} x_i = 9.94}$

جمع الـ x ات اللي بالجدول

$1-\alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \alpha/2 = 0.025$



$\therefore \Phi(Z_{\alpha/2}) = 1 - 0.025$
$= 0.975$

وبروح عالجدول بشوف هاي القيمة نتيجة إيش ؟ بتع Z أكبر

كل هاد بدنا إياه ناقص هاي اللي هي $z_{\alpha/2}$

$\therefore \boxed{Z_{\alpha/2} = 1.96}$

$\therefore$ Every thing in the equation is with it's sign!

$$P\left(9.94 - \frac{1.96(2)}{\sqrt{10}} \le \mu_x \le 9.94 + \frac{1.96(2)}{\sqrt{10}}\right) = 95\%$$

$$\boxed{P(870 \le \mu_x \le 11.1796) = 0.95}$$

* Note ①: ممكن يسألني عن الـ Confidence Interval الخاصة بالـ Errors ~~...~~ وهي بنقل $\hat{\mu}_x$ عند الـ μ_x بالنص وخلص القصة ما كانت :

$$P\left(-Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n}} \le \mu_x - \hat{\mu}_x \le Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n}}\right) = 1-\alpha$$

وإذا بدي أجاوب بمعطيات ه بطلع مع الحل ه يعني بيطل معطاة أشوي عم $\hat{\mu}_x$

$$P\left(-1.96 * \sqrt{\frac{4}{10}} \le \mu_x - \hat{\mu}_x \le 1.96 * \sqrt{\frac{4}{10}}\right) = 0.95$$

Note ② بنلاحظ بـ note [1] إنو الـ LHS و الـ RHS هم نفس الاشي ولكن تتغير إشارتهما ، وهذا يدل على القيمة المطلقة ، إذا" برضو ممكن يسألني عن ~~القيمة~~ الـ (Absolute value for the error) ، بمعنى :-

$$P\left(|\mu_x - \hat{\mu}_x| \leq Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n}}\right) = 1-\alpha$$

For ⑨ → $P\left(|\mu_x - \hat{\mu}_x| \leq 1.96 * \sqrt{\frac{4}{10}}\right) = 0.95$

Note ③ Find the value of n such that the absolute value of the error in the mean [between the true mean and the sample mean] is less than 0.1, with 95% confidence coefficient where $\sigma_x^2 = 4$.

$$P\left(|\mu_x - \hat{\mu}_x| \leq Z_{\alpha/2}\sqrt{\frac{\sigma_x^2}{n}}\right) = 0.95$$

$1-\alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \alpha/2 = 0.025 \Rightarrow \Phi(Z_{\alpha/2}) = 1-0.025$
$= 0.975$
$\therefore Z_{\alpha/2} = 1.96$ ← من الجدول نحوّل

$\therefore P\left(|\mu_x - \hat{\mu}_x| \leq 1.96 * \frac{2}{\sqrt{n}}\right) = 0.95$

$\Rightarrow 1.96 * \frac{2}{\sqrt{n}} < \frac{0.1}{1} \Rightarrow \left(\frac{1.96 * 2}{0.1}\right)^2 < (\sqrt{n})^2$

$\therefore \boxed{n = 1537}$

ملاحظة: كلما زادت نسبة التوقع اللي هي $\boxed{1-\alpha}$ ، كلما زادت الـ interval بمعنى ، زي لما الأستاذ يحكي كم بتتوقعوا بنسبة 90% العلامات بتاعت الطلاب بآخر امتحان ، فاحنا رح نحكي 10-25 ، 15-20 ، 15-25 . أما إذا قال شو بنسبة 99% تكون علامات الطلاب ، هون صار لازم أشمل كل الطلاب ونسبة الخطأ تكون بس 1% ، فمنحكي المعظم رح يكونوا يحوزوا من 5-25 هاي علامات الطلاب رح تكون بنسبة 99% ، ونسبة قليلة تجيب علامة أقل من 5 .

2 Confidence Interval on the Mean: (Variance Unknown)

$$P\left\{\hat{\mu}_x - t_{\alpha/2}\sqrt{\frac{\hat{\sigma}_x^2}{n}} \le \mu_x < \hat{\mu}_x + t_{\alpha/2}\sqrt{\frac{\hat{\sigma}_x^2}{n}}\right\} = 1-\alpha$$

على نفس سؤال قبل 6 لبس على اعتبار بعرفش ال Varince

$\boxed{\hat{\mu}_x = \frac{1}{n}\sum x_i = 9.94}$ $\quad \hat{\sigma}_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu_x)^2 = 6.51$

Number of degrees of freedom $= n-1 = 10-1 = 9$.

$1-\alpha = 0.95 \Rightarrow \boxed{\alpha = 0.05} \rightarrow \therefore \boxed{\alpha/2 = 0.025}$

$\therefore$ From the table of t-distribution:-

$t_{(\alpha/2),dof} = t_{0.025,9} = 2.263$

$$\therefore P\left\{\hat{\mu}_x - t_{\alpha/2,n-1}\frac{\hat{\sigma}_x}{\sqrt{n}} \le \mu_x \le \hat{\mu}_x + t_{\alpha/2,n-1}\frac{\hat{\sigma}_x}{\sqrt{n}}\right\} = 1-\alpha$$

$$\therefore P\left\{9.94 - 2.263\frac{\sqrt{6.51}}{\sqrt{10}} \le \mu_x \le 9.94 + 2.263\frac{\sqrt{6.51}}{\sqrt{10}}\right\} = 0.95$$

$$\therefore P\{8.11 \le \mu_x \le 11.77\} = 0.95$$

//

تم بحمد الله