4/13/16

CHOOSING THE SIZE OF THE SAMPLE

Let's define

$\therefore E = Z_{\frac{\alpha}{2}} \frac{\sigma_x}{\sqrt{n}}$

E = maximum tolerable error
= maximum amount by which the estimate of the population parameter differs from the true value of the population parameter for a specified confidence level

ex/ ESTIMATING AVERAGE AGE OF A POPULATION

We want to find a sample size so that there's a 95% chance that the average age estimated from our sample is no more than 1.5 years from the true population mean age.

For example: this means that we seek a sample size n such that if our sample mean $\bar{x}$ is 23.5, then the true mean will be between 22 and 25 years with prob. 95%

Note that we are really asking to find the sample size n, such that our 95% confidence interval for the mean will have length equal to 3.0.

In this problem, E = 1.5.

E = 0.5 years

$Z_{\frac{\alpha}{2}} = Z_{.025} = 1.96$

$\sigma_x = 2$

Find n

$E = Z_{\frac{\alpha}{2}}\left(\frac{\sigma_x}{\sqrt{n}}\right)$ * substitute in values *

$0.5 = (1.96)\frac{2}{\sqrt{n}}$ * solve for n *

4/13/16

70

ex/ ESTIMATING PERCENT OF POPULATION WEARING GLASSES 8.

We want to find a sample size so that there's a 90% chance that the percentage of people wearing eyeglasses in our sample is within 2 percentage points of the true population percentage.

For example: this means that we seek a sample size n such that if our sample percentage of people wearing eyeglasses is 28%, then it is 90% certain that the percentage of people in the population who wear eyeglasses is between 26% and 30%.

Note that we are really asking to find the sample size n such that our 90% confidence interval for the proportion will have length equal to .04.

In this problem $E = .02$

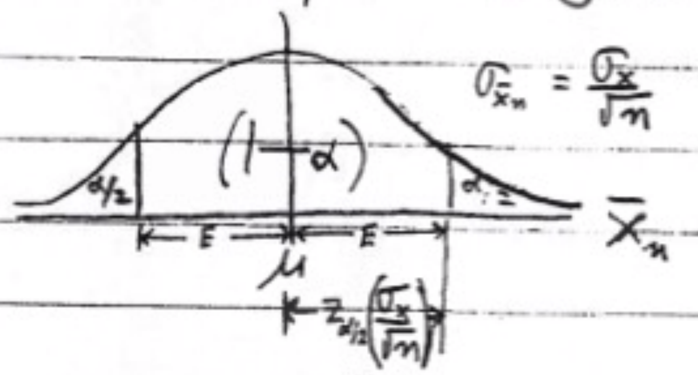
*solve the same way you solved for previous problem**

4/13/16

FINDING SAMPLE SIZE WHEN ESTIMATING A POP. MEAN μ

Look at the probability distribution of $\bar{X}_n$, the mean of a sample of size n

$\sigma_{\bar{x}_n} = \frac{\sigma_x}{\sqrt{n}}$



From confidence interval computation

$\therefore \; E = Z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}}$ and we know everything except n, so we solve for n

note: we ... need in ... to do this

ex/ How many students do we have to sample to be 95% sure that our point estimate of the average size is correct within 0.5 years? Assume we know that σ_x is about 2 years.

$E = .5$ years
$Z_{\alpha/2} = Z_{.025} = 1.96$
$\sigma_x = 2$
$n = ?$

when you don't know σ_x, can use range/4 for σ_x if not too skewed p 229

So $\therefore \; E = Z_{\alpha/2} \frac{\sigma_x}{\sqrt{n}}$

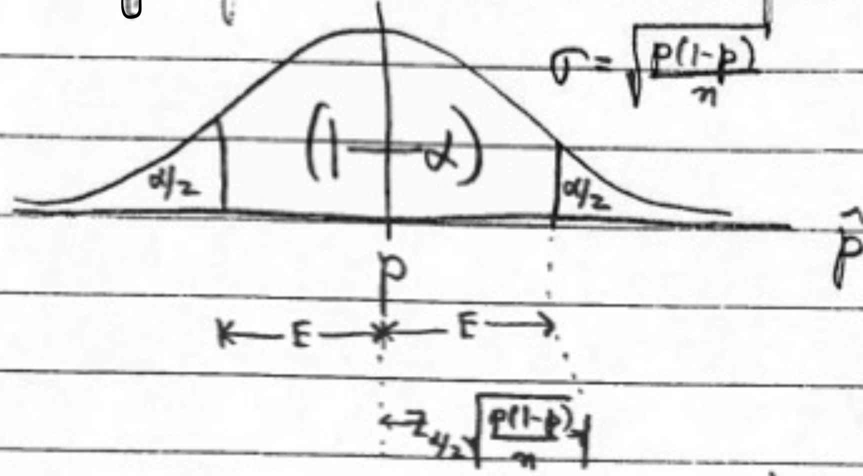
$.5 = 1.96 \frac{2}{\sqrt{n}}$

$\therefore \sqrt{n} = \frac{(1.96)(2)}{.5} = 7.84$

$n = 61.46$: so we need a sample size of $\underline{\underline{62}}$

FINDING SAMPLE SIZE WHEN ESTIMATING A POPULATION PROPORTION

Look at the probability distribution of $\hat{p}$, the sample proportion in a sample of size n:



From confidence interval computation

So: $E = z_{\alpha/2}\left(\sqrt{\frac{p(1-p)}{n}}\right)$

and we can solve this for n. If p is unknown, the conservative approach is to use $p=.5$, since $\sqrt{\frac{p(1-p)}{n}}$ is at its largest when $p=.5$

ex/ We wish to estimate the proportion of people who wear glasses, and we want a sample size large enough so that our estimate $\hat{p}$ (sample proportion) is 90% likely to be within 1 percentage point of the true proportion. Assume we know that the proportion is about 1/3 (else we would use $\frac{1}{2}$)

$E = .01 \quad p = .33 \quad z_{.05} = 1.64$

$$E = z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

$$.01 = 1.64\sqrt{\frac{.33(.67)}{n}}$$

square both sides:

$$.0001 = 2.6896\left(\frac{(.33)(.67)}{n}\right)$$

$$n = \frac{2.6896(.33)(.67)}{.0001} = \underline{5947}$$

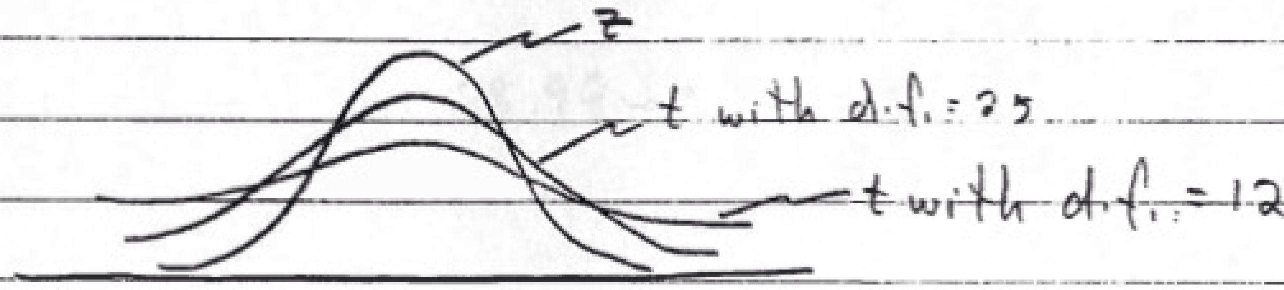
NOTE: IF $n = 1000$, $E = .024$ or 2.4 percentage points

TO GO FROM $E = .024$ TO $E = .01$ REQUIRES ~SIX TIMES THE SAMPLE SIZE

4/13/16

SMALL-SAMPLE ESTIMATION

When you want an interval estimate (confidence interval) for the population mean μ, you must use the t-distribution instead of the z-distribution if σ_x is unknown (must be estimated by s_x) and the sample size is less than 30.



The t-distribution changes with sample size n and is characterized by "degrees of freedom" (d.f.), where d.f. = $n-1$

smaller $n \Rightarrow$ more variable (t depends on 2 r.v.'s: $\bar{x}$ and s)

For $n \geq 30$, the t-distribution is virtually identical to the z-distribution.

So our $(1-\alpha)\%$ interval estimate (confidence interval) is

$$\bar{x} \pm t_{n-1,\alpha/2}\left(\frac{s_x}{\sqrt{n}}\right)$$

4/13/16

89

ex/ Find a 98% confidence interval for μ if $\bar{x} = 28.3$, $s_x = 1.22$, and the sample size is 20

$$\bar{x} \pm t_{n-1,\alpha/2}\left(\frac{s_x}{\sqrt{n}}\right) = 28.3 \pm t_{19,.01}\left(\frac{1.22}{\sqrt{20}}\right)$$

$z = 2.33$ for $\alpha = .01$

$$= 28.3 \pm 2.539\left(\frac{1.22}{\sqrt{20}}\right)$$

$$= (27.60, 28.99)$$

OR $P(27.60 \le \mu \le 28.99) = .98$

NOTE: small samples are not used to estimate population proportions because the interval estimates are so wide.

TO ESTIMATE (TEST) POPULATION MEAN

FLOW CHART