Confidence Interval - Assignment

1. Given, sample size (n) = 1000, $\mu = 1000000$

avg wt of man in our sample = 180 pounds
$\bar{x} = 180$

Standard deviation $\sigma = 30$

standard error of mean (SE) $= \frac{\sigma}{\sqrt{n}} = \frac{30}{\sqrt{1000}} = 0.96$

Confidence limits for 95% confidence limit

is $= \mu_s \pm \underbrace{2}_{\text{z value for 95\% confidence interval}} SE$

$180 \pm 2(0.96)$

C. Limits $= 180 \pm 1.92$

2. Given, $\sigma = 3.6$ minutes

a) $n = 120$, $\bar{x} = 16.2$ minutes, 92% confidence interval = ?

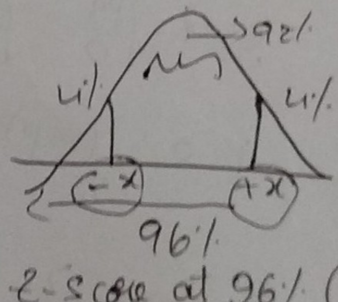
Standard error of mean $= \frac{\sigma}{\sqrt{n}} = \frac{3.6}{\sqrt{120}} = \frac{3.6}{10.95}$
$= 0.328$

Confidence intervals
$= \mu \pm SE \times z\text{-score}$

$= 16.2 \pm 1.75 \times (0.328)$

15.626, 16.774

= 15.626 − 16.774 (Limits)



z-score at 96% (0.96)

we need (−x) to (+x)
that is 96%

import scipy.stats as stats
stats.norm.ppf(0.96)

z-score at area 0.96
= 1.75

b) $n = ?$, ± 15 seconds $= 0.25$ minute

confidence 92%, same $z = 1.75$

Standard error $\left(\sqrt{\frac{pq}{n}}\right) \frac{\sigma}{\sqrt{n}}$, $z = 1.75$

we have margin of error $= z * SE$

$$0.25 = 1.75 \times \frac{3.6}{\sqrt{n}}$$

$$\sqrt{n} = \frac{3.6 \times 1.75}{0.25}$$

$$\frac{6.3}{0.25}$$

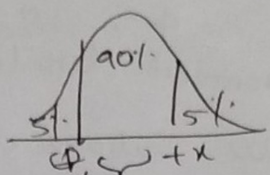
$\sqrt{n} = 25$, $n = 5.01$

→ 5 workers should be involved in this study inorder to have mean assembly time estimated upto ± 15 seconds with 92% confidence.

③ Given, (a) margin error = 2% (0.02), confidence = 90%

$n = ?$ Standard Error of sample proportion $z = \sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$

z-Scale



$\alpha < 95\%$

z-Score = 1.6845

$\hat{P}$ → probability in sample;

assuming here $P = 0.5$

$ME = Z * SE$

$0.02 = 1.645 \times \sqrt{\frac{0.5 \times 0.5}{n}}$

$\frac{0.02}{1.645} = \sqrt{\frac{0.25}{n}}$

$0.0121 = \sqrt{\frac{0.25}{n}}$ Squaring

$0.000146 = \frac{0.25}{n}$

$\boxed{n = 1712}$

(b) Given, $n = 1000$ (sample size)

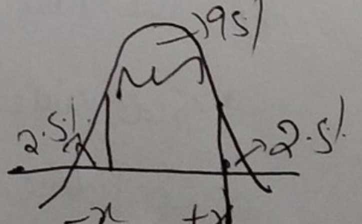
$P = 400,\ (1-P) = 600$

$P = 0.4,\ (1-P) = 0.6$

↓

(customers happy with their purchase)

95% confidence Interval for 'p'



97.5%

Confidence Level $= 0.4 \pm 1.96 \left(\sqrt{\frac{(0.4)(0.6)}{1000}} \right)$ z-score = 1.959 ≃ 1.96

$= 0.4 \pm 1.96 \left(\sqrt{\frac{0.24}{1000}} \right)$

$= 0.4 \pm 1.96 (0.015)$

$= 0.4 \pm 0.0294$

Confidence limits = 0.4294, 0.371

(4) Given, Repeatedly weighed 4 times, measurement of weight

$n = 4$, 0.95, 1.02, 1.01, 0.98

Normal distribution with mean μ, σ unknown

a) Use above data to compute 95% confidence interval μ.

μ of sample, $n=4$, [0.95, 1.02, 1.01, 0.98]

$= (0.95+1.02+1.01+0.98)/4$

$= 3.96/4 = 0.99$

Standard deviation $s = \sqrt{(0.99-0.95)^2 + (0.99-1.02)^2 + (0.99-1.01)^2 + (0.99-0.98)^2 / 4}$

$= \sqrt{(0.04)^2 + (-0.03)^2 + (-0.02)^2 + (0.01)^2 / 4}$

$= \sqrt{0.0016 + 0.0009 + 0.0004 + 0.0001 / 4}$

$= \sqrt{0.003/4} = \sqrt{0.00075} = 0.02738$

$= \simeq 0.0274$

Sample is small, standard deviation of population is not known, Pop follows Normal distribution

So, we take t-distribution, degree of freedom $= n-1 = 4-1 = 3$.

t-score value for 95% confidence is $t = 3.182$

Confidence levels = Sample mean $\pm t *$ $\frac{\text{Standard deviation of sample}}{\sqrt{n}}$

$= 0.99 \pm \frac{3.182(0.0274)}{2}$

$= 0.99 \pm 3.182(0.0137)$

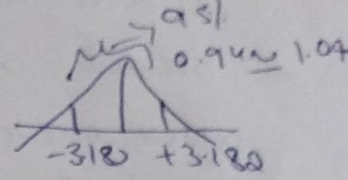
$= 0.99 \pm 0.0435$

$(0.9465, 1.033)$

4 (b) Do these data give evidence at 5% significance level that the scale is not accurate?

t-score at 3 df at 0.025

= ~~2.353363~~ 3.182



Data lies within 95% values,

No evidence at 5% significance level that scale is inaccurate.

5 Given, $\mu = 45$ seconds

$n = 9$ (sample), Sample mean $\bar{x} = 49.2$ seconds

Sample Standard deviation = 3.5

Null Hypothesis ⇒ College students complete maze in 45 seconds

Alternate Hypothesis ⇒ College Students cannot complete maze in 45 / either may be greater than 45 or less than 45 seconds

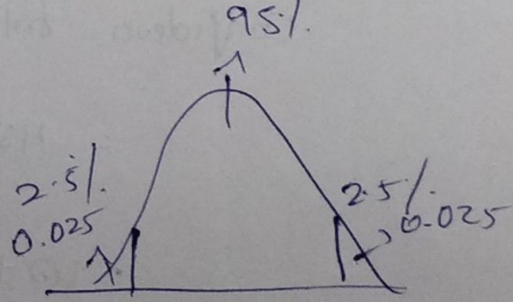
($n < 30$, $9 < 30$) Population standard deviation unknown, we use t-distribution

$$t\text{-score} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

$$= \frac{49.2 - 45}{\frac{3.5}{\sqrt{9}}}$$

$$= \frac{4.2}{\frac{3.5}{3}} = \frac{4.2}{1.166}$$

$$= 3.6002$$



5% significance level

P-value at t-score (3.6002) is 0.005

df, t-8 degree of freedom for 5% significance (0.05) (left tail-0.025) right tail 0.025)

0.025 t-score is 2.306

Here, for sample we got t-score 3.6 > 2.3
Outside tail limits, so, we reject Null Hypothesis
& Students doing maze in 45 seconds not true,
after exercising, their mean time has changed
with 5% significance level.

6. Given, $n = 64$, sample mean = 42 min
Standard deviation = 5 min

$$\text{Standard error} = \frac{\sigma}{\sqrt{n}}$$

$$= \frac{5}{\sqrt{64}} = 5/8 = 0.625$$

$n > 30$, population S.D. known, we calculate

$$z\text{-score} = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

$$= \text{at } 95\% \text{ confidence} = 1.96$$

$$\text{Confidence Interval} = \bar{x} \pm z\text{-score} \times SE$$

$$= 42 \pm 1.96\left(\frac{5}{8}\right)$$

$$= 42 \pm 1.96(0.625)$$

$$= [42 + 1.96(0.625),\ 42 - 1.96(0.625)]$$

$$[42 + 1.225,\ 42 - 1.225]$$

$$= [43.225,\ 40.775]$$

Mean Installation Time Confidence Interval

$$= [40.775,\ 43.225]$$

⑦ Given, $n = 17$ $\Sigma d_i = -3.50$ $\Sigma d_i^2 = 19.13$

90% confidence for avg diff (right - left)

$\bar{x} = \frac{\Sigma d_i}{n} = \frac{-3.50}{17} = -0.205$

variance $\frac{\Sigma d_i^2}{n} = \frac{19.13}{17} = 1.125$

$\sigma = 1.06$

z-score for 90% = 1.645

Confidence Interval $= -0.205 \pm 1.645(1.06)$

$= \left[-0.205 - 1.645(1.06), -0.205 + 1.645(1.06)\right]$

$= [-0.205 - 1.7437, -0.205 + 1.7437]$

$= [-1.9487, 1.5387]$

⑧ Given, $n = ?$, 95% confidence interval for mean = ?

interval no longer than 1cm,

Population Normal with variance 9^2

$\sigma = 9$

z-score = 1.96 (for 95%)

Smallest sample size = ?

< 1cm (interval)

$$\bar{x} \pm \underbrace{\frac{(1.96)(9)}{\sqrt{n}}}_{} < 1$$

$$\frac{1.96(9)}{\sqrt{n}} < 1$$

$17.64 < \sqrt{n}$

$= 311.1696 < n$

$n > 311.1696$ (Smallest 'n' could be 311.1696)

(9) Given, $\mu = 150\$$, sample size$(n) = 16$

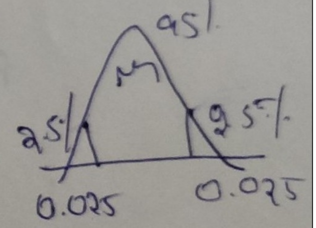
Sample Mean$(\bar{x}) = \$141$, Sample standard deviation $(\sigma_s) = 4$

$n < 30$, population follows Normal distribution, pop standard deviation unknown

So, we take t-score

t-score at df = 15, 95% confidence interval

is $t = 2.1314$



Confidence Interval $= \bar{x} \pm t \cdot \dfrac{\text{Standard deviation of sample}}{\sqrt{n} \text{ (sample size)}}$

$= 141 \pm 2.1314 \left(\dfrac{4}{\sqrt{16}}\right)$

$= (141 - 2.1314,\ 141 + 2.1314)$

$= [138.8686, 143.1314]$

(10) Given, Sample size(n) = 17096

frequent binge drinkers = 3314

Proportion $p = \frac{3314}{17096}$

90% confidence, Z-score = 1.645

Confidence interval = $p \pm z.\sqrt{\frac{pq}{n}}$

$$= \frac{3314}{17096} \pm 1.645\left(\sqrt{\frac{(3314)(17096-3314)}{17096}}\right)$$

$$= 0.1938 \pm 1.645\left(\sqrt{\frac{(3314)(13782)}{17096}}\right)$$

$$= 0.1938 \pm 1.645\left(\sqrt{2671.59}\right)$$

$$= 0.1938 \pm 1.645(51.68)$$

$$= 0.1938 \pm 85.0136$$

$$(85.2074, 84.8198)$$

(11) Given, n (sample size) = 100, Sample mean $\bar{x} = 49$,

pop standard deviation (σ) = 4.49

90% confidence interval to estimate pop mean, Z-score = 1.645

$$\bar{x} \pm z\text{-score} \times SE$$

$$= 49 \pm 1.645\left(\frac{4.49}{\sqrt{100}}\right)$$

$$= 49 \pm 1.645(0.449)$$

$$= 49 \pm (0.7386)$$

Confidence interval = (48.2614, 49.7386)

(12) Given, Sample size $(n) = 1200$ clicks

proportion of fraudulent clicks $p = \frac{175}{1200}$

95% confidence level

$$p \pm z\text{-score}\sqrt{\frac{pq}{n}}$$

z-score for 95% confidence $= 1.96$

$$\frac{175}{1200} \pm 1.96\left(\sqrt{\frac{(175)(1200-175)}{1200}}\right)$$

$$= 0.1458 \pm 1.96\left(\sqrt{\frac{(175)(1025)}{1200}}\right)$$

$$= 0.1458 \pm 1.96(12.226)$$

$$= 0.1458 \pm 23.96$$

proportion of fraudulent clicks $= (23.8142, 24.1058)$

confidence interval

(13) Given, $n = 59$, $p =$ left-handed player $= \frac{15}{59}$

95% confidence True percentage = ?

$$\frac{15}{59} \pm (1.96)\left(\sqrt{\frac{15 \times 44}{59}}\right)$$

$$= 0.254 \pm 1.96(\sqrt{11.18})$$

$$= 0.254 \pm 1.96(3.343)$$

$$= 0.254 \pm 6.55228$$

$$= 6.296, 6.806$$

(14) Given, margin of error ≤ 100\$

90% confidence, z-score = 1.645

Sample size (n) = ?, standard deviation (σ) = 475

$$ME = Z \times \frac{\sigma}{\sqrt{n}}$$

$$100 = 1.645 \times \frac{475}{\sqrt{n}}$$

$$\sqrt{n} = \frac{1.645 \times 475}{100}$$

$$\sqrt{n} = 7.81375$$

$n = 61.05 \rightarrow$ Sample size they can take

for no of chain of food restaurants with given confidence & standard deviations is 61.05.

(15) Given for 5-minute interval, the arrival count of customers is 68, 42, 51, 57, 56, 80, 45, 39, 36, 79

$n = 10$

95% confidence interval to estimate mean value for all 5-minute intervals; Here ($n < 10$), so we take t-score

t-score at 0.025 with $df = 9 (10-1)$

$= 2.26$

$$\bar{X} = \frac{68+42+51+57+56+80+45+39+36+79}{10}$$

$$\bar{X} = 55.3$$

Standard deviation $= \sqrt{\sigma^2}$

$$\sigma^2 = (68-55.3)^2 + (42-55.3)^2 + (55.3-51)^2 + (55.3-57)^2 + (55.3-56)^2 + (55.3-80)^2 + (55.3-45)^2 + (55.3-39)^2 + (55.3-36)^2 + (55.3-79)^2 / 10$$

$$= (12.7)^2 + (13.3)^2 + (4.3)^2 + (-0.7)^2 + (-24.7)^2 + (10.3)^2 + (16.3)^2 + (19.3)^2 + (-23.7)^2 / 10$$

$$= 161.29 + 176.89 + 18.49 + 0.49 + 610.09 + 106.09 + 265.69 + 372.49 + 561.69 / 10$$

$$\sigma^2 = 2273.21/10 \Rightarrow \sigma = \sqrt{\frac{2273.21}{10}} = 15.077$$

Confidence interval

$$= \bar{x} \pm t \cdot \frac{\sigma}{\sqrt{n}}$$

$$= 55.3 \pm 2.26 \left(\frac{15.077}{\sqrt{10}}\right)$$

$$= 55.3 \pm 2.26 (4.768)$$

$$= (55.3 - 10.77,\ 55.3 + 10.77)$$

$$= [44.53, 66.07]$$