Confidence interval w/ C confidence level ... March 8th

... $\underline{\underline{C}}$ ... probab. that the "actual" mean μ falls within your confidence interval.

. $\mathbb{P}$[two independent samples both result in a confidence interval containing μ] = C^2

· $\mathbb{P}$[at least one of two independent samples results in a confidence interval containing μ] =

= 1 - $\mathbb{P}$[<u>none</u> of the two independent samples results in a confidence interval containing μ] =

$$= 1-(1-C)^2$$

$$= (1-1+C)(1+1-C) = C(2-C) =$$

$$= 2C - C^2 = C + C - C^2$$

UNIVERSITY OF TEXAS AT AUSTIN

Problem Set # 6

The t-procedure.

Provide your **complete solution** for the following problems.

Problem 6.1. When examining data to determine if a $t-$procedure can be used, which of the following are useful?

(a) Histogram
(b) $Q-Q$ plot
(c) Boxplot
(d) All of the above.
(e) None of the above.

Problem 6.2. A simple random sample of 20 third-grade children from a certain school district is selected. Each child is given a test to measure his/her reading ability. You are interested in calculating a 95% confidence interval for the population mean score based on this SRS. The sample mean score is 64 points and the sample standard deviation is 12 points. What is the margin of error associated with the 95% confidence interval?

Problem 6.3. The hypotheses $H_0 : \mu = 10$ versus $H_a : \mu \neq 10$ are examined using a sample of size $n = 18$. The one-sample $t-$statistic has the value of $t = -2.05$. Between what two values does the P-value of this test fall?

(a) $0.01 < P-$value < 0.02,
(b) $0.02 < P-$value < 0.025,
(c) $0.025 < P-$value < 0.05,
(d) $0.05 < P-$value < 0.10
(e) None of the above.

Problem 6.4. The time (in seconds) needed to clear a certain *Super Mario* run is normally distributed with an unknown mean μ. The local *Bowser basher* league reports the mean clearance time of at most 75 seconds. You are incredulous of their statements and decide to test your hypothesis.

i. What are your null and alternative hypotheses?
ii. Note that the standard deviation of the clearance time is not specified. Which test statistic are you going to use to test your hypotheses?
iii. Let the sample size be 20. What is the distribution of your test statistic?
iv. You gather the data, calculate $\bar{x}$ and s and evaluate the $t-$statistic. The value you obtain equals 1.68. Is your null hypotheses rejected at the siginificance level 0.05?

Problem 6.5. When testing the hypothesis of equal means of two independent random samples, what distribution assumption is necessary for the randomly sampled data?

Problem 6.6. When testing the hypothesis of equal means of two independent random samples, which distribution is used?

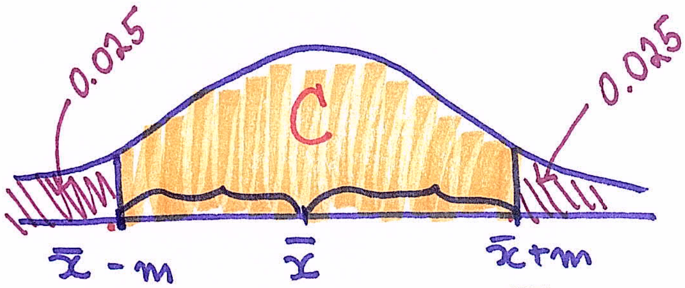
#6.2. $\boxed{n = 20}$

$C = 0.95$

Conf. inteval : $\bar{x} \pm m$

margin of error

Use the t-procedure since σ not given.



t - statistic $t = \dfrac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \sim t(df = 19)$

df... degrees of freedom

$df = n - 1$

margin of error: $m = \pm t^* \dfrac{s}{\sqrt{n}}$

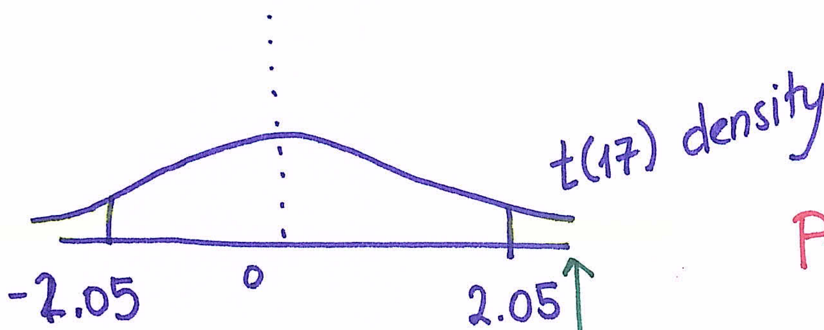
$t^* \ldots \begin{cases} \text{based on } df \\ \text{and } C \end{cases}$

$t^* = 2.093 \quad \Rightarrow \quad m = 2.093 \cdot \dfrac{12}{\sqrt{20}} \approx 5.616.$

b.

6.3. TS… test statistic

$t = -2.05$ w/ $df = 17$



P-value: area of the yellow region

From the t-table

$\Rightarrow$ upper-tail probab. is between 2.5% and 5%

$\Rightarrow$ $0.05 < \text{P-value} < 0.10$ $\Rightarrow$ (d).

6.4. (i.) $H_0: \boxed{\mu = 75}$

$H_a: \mu > 75$ μ... the unknown population mean

(ii) TS... test statistic

$$t = \frac{\bar{x} - 75}{s/\sqrt{n}} \quad \text{w/} \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

(iii.) sample size 20 $\Rightarrow$ $t(df = 19)$

(iv.) We get $t = 1.68$ from our sample. With $\alpha = 0.05$

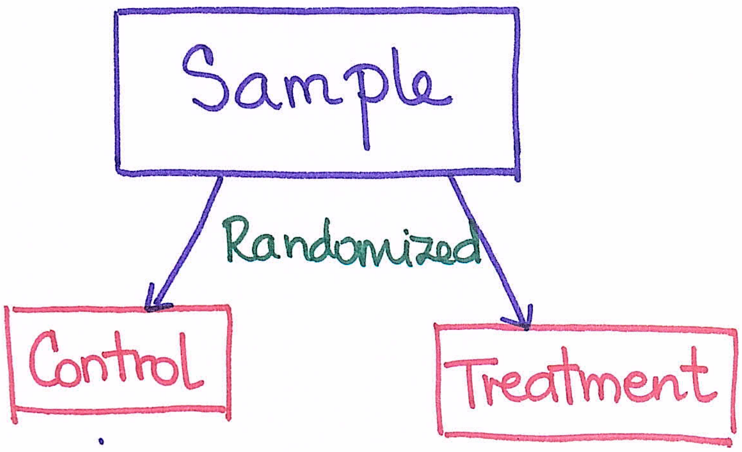
$\Downarrow$

P-value > 0.05 $\Longrightarrow$ Fail to reject.

Example. [Drug Testing]

"Placebo/ sugar pill" vs. "New drug"

"old drug" vs. "New drug"



μ_1 ... mean of the population 1 model

μ_2 ... mean of the population 2 model

Usually: $H_0: \mu_1 = \mu_2$ (no effect)

$H_a: \mu_1 > \mu_2$ or $\mu_2 > \mu_1$ or $\mu_1 \neq \mu_2$

Interested $\mu_1 - \mu_2$

$\Rightarrow$ take a closer look @ the ESTIMATOR of this difference:

$$\bar{X}_1 - \bar{X}_2 \quad \ldots \text{ rnd variable}$$

Assumption:

- the population distributions are normal;
- $\bar{X}_1$ and $\bar{X}_2$ are independent

$$\Rightarrow \bar{X}_1 \sim N\left(\text{mean} = \mu_1, \text{var} = \frac{\sigma_1^2}{n_1}\right)$$

$$\bar{X}_2 \sim N\left(\text{mean} = \mu_2, \text{var} = \frac{\sigma_2^2}{n_2}\right)$$

$$\Rightarrow \bar{X}_1 - \bar{X}_2 \sim N\left(\text{mean} = \mu_1 - \mu_2, \text{var} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_1^2}{n_2}\right)$$

H_0: $\mu_1 = \mu_2 \iff (\mu_1 - \mu_2) = 0 = (\mu_1 - \mu_2)_0$

Under the null:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - \underbrace{\cancel{(\mu_1 - \mu_2)_0}}_{=0}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

8