Today, we will:

- Do a review example problem – *t* PDF and chi-squared PDF
- Review the pdf module: **Correlation and Trends** and do some example problems
- Review the pdf module: **Regression Analysis** and do some example problems

Example: Estimating population mean and population standard deviation

Given: 20 ball bearings are pulled from the assembly line, and their diameters are measured. The sample mean is 2.56 mm and the sample standard deviation is 0.240 mm.

(*a*) To do: Estimate the population mean and its confidence interval for 98% confidence level.

(*b*) To do: Estimate the population standard deviation and its confidence interval for 98% confidence level.

Solution: · Our best estimates are $\mu \approx \bar{x} = 2.56$ mm

$\sigma \approx S = 0.240$ mm

· But, how confident are we in these estimates? — We want to know the range of μ & σ for 98% confidence level.

(a) · $n = 20 \rightarrow df = 20 - 1 = \underline{19 = df}$ & $\boxed{\text{Use } t \text{ PDF}}$

· @ 98% confidence, $C = 0.98$, $\alpha = 1 - C = 0.02$, $\alpha/2 = 0.01$

· Table @ 98% & $df = 19 \rightarrow \underline{t_{\alpha/2} = 2.5395}$

OR, using Excel, $= \text{TINV}(0.02, 19)$ (α, df)

· So, $\boxed{\mu = \bar{x} \pm t_{\alpha/2} \frac{S}{\sqrt{n}}} \rightarrow \boxed{\mu = 2.56 \pm 0.136 \text{ mm to 98\% confidence}}$ ANSWER

(b) · $n = 20$, df is still $n - 1 = 19$ in this case. $\boxed{\text{Use } \chi^2 \text{ PDF}}$

· recall, @ 98% confidence, $C = 0.98$

$\alpha = 0.02 \rightarrow \frac{\alpha}{2} = 0.01$, $1 - \frac{\alpha}{2} = 0.99$

[area of one tail] [area of everything but one tail]

· @ 98% confidence, see Table for critical χ^2 values:

$\chi^2_{1-\alpha/2} = \underline{7.6327}$ [OR, Excel $= \text{CHIINV}(0.99, 19)$] $1 - \alpha/2$, df

$\chi^2_{\alpha/2} = \underline{36.1908}$ [OR, Excel $= \text{CHIINV}(0.01, 19)$] $\alpha/2$

Thus, $$\sigma_{low} = \sqrt{\frac{(n-1)\ S^2}{\chi^2_{\alpha/2}}} = 0.174 \text{ mm}$$

$$\sigma_{high} = \sqrt{\frac{(n-1)\ S^2}{\chi^2_{1-\alpha/2}}} = 0.379 \text{ mm}$$

Final estimate:

$$\boxed{0.174 \le \sigma \le 0.379 \text{ mm} \quad \text{to 98\% confidence}}$$

Comments:

1) There is a 2% chance that σ actually lies <u>outside</u> of this range (both tails)

2) There is a 1% chance (one tail) that σ is <u>smaller than</u> 0.174 mm. If so, this would be a <u>good</u> thing! (means that the manufacturing process is very consistent)

3) There is also a 1% chance (one tail — the other one) that σ is <u>larger than</u> 0.379 mm.

If so, this would be <u>bad</u> → means that there are some inconsistencies in the manufacturing process.

In general, we are concerned with σ being too <u>big</u>, not with σ being too <u>small</u>.

Example: Estimating population mean from a sample

Given: A company produces resistors by the thousands, and Mark is in charge of quality control. He picks 20 resistors at random as a sample, and calculates the sample mean $\overline{x} = 8.240$ kΩ and the sample standard deviation $S = 0.314$ kΩ.

To do: Estimate the **confidence interval of the population mean** (as a ± value) to standard 95% confidence level, i.e. fill in the blank: $\mu = 8.240 \;\pm$ __________ kΩ.

Give your answer to three significant digits.

A portion of the critical *t* table is shown here for convenience.

Values of $t_{\alpha/2}$ (critical values) for the student's *t* distribution

	90% confidence	95% confidence	98% confidence	99% confidence
α = →	0.10	0.05	0.02	0.01
df = ↓				
1	6.3137	12.7062	31.8210	63.6559
2	2.9200	4.3027	6.9645	9.9250
3	2.3534	3.1824	4.5407	5.8408
16	1.7459	2.1199	2.5835	2.9208
17	1.7396	2.1098	2.5669	2.8982
18	1.7341	2.1009	2.5524	2.8784
19	1.7291	2.0930	2.5395	2.8609
20	1.7247	2.0860	2.5280	2.8453
21	1.7207	2.0796	2.5176	2.8314

Solution:

$t_{\alpha/2}$ @ 95% confidence & df = 19

- Use t PDF
- $\mu = \bar{x} \pm t_{\alpha/2} \dfrac{S}{\sqrt{n}}$
- Here, df = n−1 = 20−1 = 19 = df Table: $t_{\alpha/2} = 2.0930$
- Plug in numbers: $\mu = 8.240 \pm (2.0930) \dfrac{0.314}{\sqrt{20}}$ kΩ

Caution: Use n=20 here, not df = 19

Answer: $\boxed{\mu = 8.240 \pm 0.147 \text{ k}\Omega \text{ to } 95\% \text{ confidence}}$

Example: Linear correlation coefficient

Given: Matt measures both the shoe size and the weight of 18 football players. He performs a linear regression analysis of shoe size (y variable) as a function of weight (x variable). He calculates $r_{xy} = 0.582$.

To do: To what confidence level can Matt state that a football player's shoe size is correlated with his weight?

Solution:

· At $n = 18$, look up critical linear correlation coefficient and interpolate:

C (confidence level)	r_t (critical lin. corr. coeff.)	
98%	0.54255	
$\boxed{}$	0.582	← our r_{xy}
99%	0.58971	

Interpolate to get C

$\boxed{\text{ANSWER: } C = 98.84\% \text{ confidence level}}$

Example: Linear correlation coefficient

Given: Several measurements are taken in a wind tunnel of pressure difference as a function of distance normal to the direction of flow over a body.

Data point	x (mm)	ΔP (kPa)
1	0	0.356
2	2.0	0.679
3	4.0	0.478
4	6.0	0.564
5	8.0	0.390
6	10.0	0.581
7	15.0	0.805
8	20.0	0.723

See Example Excel file, "Class_Example_Correlation.xls"

[We could use the eq. for r_{xy} given in the learning module, but it is a lot of work to calculate this "by hand". Excel, Matlab, etc. have functions to do it easily.]

(*a*) To do: Calculate the linear correlation coefficient

(*b*) To do: To what confidence level can we state that there is a trend in the data?

Solution: (a) → See Excel, get $\boxed{r_{xy} = 0.6589}$

(b) @ $n = 8$

C	r_t	
90%	0.62149	
$\boxed{}$	0.6589	← our r_{xy}
92.5%	0.65985	

Interpolation gives $\boxed{C = 92.44\%}$

We are confident to 92.4% that there is a trend.

[Since $C < 95\%$, we do not accept the trend]

Example: Regression Analysis

Given: The same pressure vs. distance measurements of the previous problem.

To do: Perform a linear regression analysis – plot the best-fit straight line and compare the fitted curve to the data points.

Solution:

See Excel spreadsheet – I will show in class how to do the regression analysis in Excel.