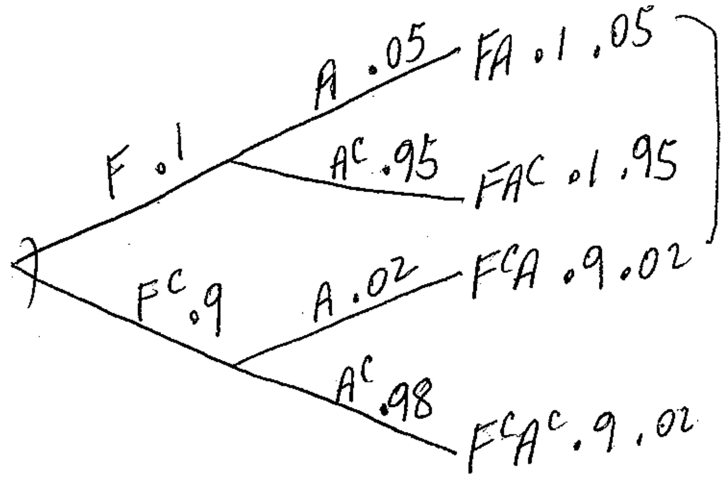
STT351-001
Final Exam

1-2. TREE.

10% of parts are flawed.
5% of flawed parts **appear flawed**
2% of non-flawed parts appear flawed.
Make a tree.



1. P(part appears flawed)

$= P(FA) + P(F^cA)$
$= .1\ .05 + .9\ .02$

2. P(part is flawed | part appears flawed) $= P(FA)/P(F)$

$$= \frac{.1\ .05}{.1\ .05 + .9\ .02}$$

3-4. CI and TEST for MEAN μ. A sample of n = 7 prescription eyeglass lenses is drawn from a process **under statistical control**. Each of these **seven** is subjected to measurements which determine an overall score x = "conformity to prescription."

3. Formula for 90% CI for μ with appropriate n, t or z score. $DF = 7-1 = 6$

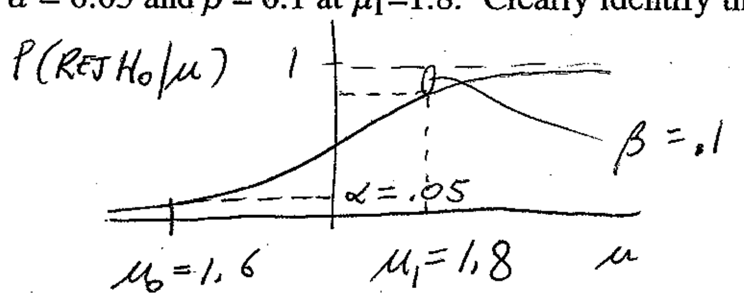
DF 90% CENTRAL
6 $\boxed{1.943}$ 90% $t_{DF=6}$ $t = 1.943$

$\bar{X} \pm 1.943 \frac{s}{\sqrt{7}}$

4. Sketch curve P(reject $H_0 \mid \mu$) vs μ for a 1-sided test of

$H_0 : \mu = 1.6$ vs $H_1 : \mu > 1.6$

with $\alpha = 0.05$ and $\beta = 0.1$ at μ_1=1.8. Clearly identify these elements in your sketch.



5-6. Probability Rules. $P(A) = 0.6$, $P(A \mid B) = \cancel{0.7}$ 0.5, $P(B) = \cancel{0.9}$ 0.8.

5. $P(A \cup B)$ $= P(A) + P(B) - P(AB)$

$= 0.6 + 0.8 - P(B)P(A|B) = 1$

NOT REQUIRED

6. $P(B \mid A)$ $= \dfrac{P(AB)}{P(A)} = \dfrac{0.8 \; 0.5}{0.6} = \dfrac{0.4}{0.6}$

7-8. Normal Table. Diameter D is normally distributed with $\mu = 0.5$ and $\sigma = 0.01$.

7. Standard score z for diameter 0.52. $z = \dfrac{d - \mu}{\sigma} = \dfrac{0.52 - 0.5}{0.01} = 2$

8. $P(Z \text{ in } [1.34, 2.27])$ $= P(Z < 2.27) - P(Z < 1.34)$

$= .9984 - .9099$

eg .04 .07

1.3 $\boxed{.9099}$

2.2 $\boxed{.9884}$

9-10. Estimates. A sample of n = 100 is selected **with-replacement** and with equal probability from a population of size 200. This sample has mean $\overline{x} = 2.1$ with $s = 0.6$.

9. Estimate the sd $\sigma_{\overline{X}}$ of sample mean $\overline{x}$ EST OF $\sigma_{\bar{x}} = \dfrac{\sigma_{pop}}{\sqrt{n}}$ IS $\dfrac{s}{\sqrt{n}} = \dfrac{0.6}{\sqrt{100}}$

10. Repeat (a) if the sample is ~~withOUT-replacement~~ " w/o "

w/o REPL $\sigma_{\bar{x}} = FPC \dfrac{\sigma_{pop}}{\sqrt{n}} = \sqrt{\dfrac{N-n}{N-1}} \dfrac{\sigma_{pop}}{\sqrt{n}}$

EST^{d} BY $\sqrt{\dfrac{N-n}{N-1}} \dfrac{s}{\sqrt{n}} = \sqrt{\dfrac{200-100}{200-1}} \dfrac{0.6}{\sqrt{100}}$

11-12. CI for difference of means. A with-replacement sample of 70 parts from supplier A finds sample mean breaking strength = 560 with sample sd s = 29. Independently of this, a with-replacement sample of 70 parts from supplier B finds sample mean breaking strength 540 with sample sd s = 38.

11. Determine 95% z-based CI for the difference $\mu_A - \mu_B$.

$$\bar{x} - \bar{y} \pm z\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}} \quad \text{i.e.} \quad (560-540) \pm 1.96\sqrt{\frac{29^2}{70} + \frac{38^2}{70}}$$

12. Determine the z-statistic needed to test the hypothesis H_0 that the mean breaking strength of parts is the same for supplier A as it is for supplier B.

$$\frac{\bar{x} - \bar{y} - 0}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}} \quad (\mu_x - \mu_y = 0 \text{ UNDER } H_0) \quad = \frac{560-540}{\sqrt{\frac{29^2}{70} + \frac{38^2}{70}}}$$

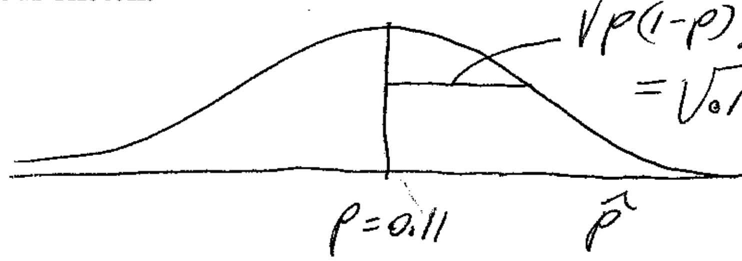
ASIDE: TEST REJECTS H_0 IF $|t_{STAT}| > t \sim z$ CRITICAL

WHERE $\alpha/2$, $1-\alpha$, $\alpha/2$; z

13-14. Binomial. 11% of a population of vases is damaged in shipment (i.e. p = 0.11). Denote by $\hat{p}$ the fraction-defective in a with-replacement sample of 100 parts from this population.

13. Sketch the normal approximation of the distribution of $\hat{p}$, identifying the numerical mean and sd of $\hat{p}$ in your sketch.

$$\sqrt{p(1-p)}/\sqrt{n} = \sqrt{.11 \cdot .89}/\sqrt{100}$$



14. Determine a 95% z-based CI for p if we find that out of a particular sample of 100 vases there are 9 damaged in shipment.

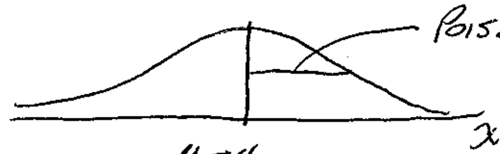
$$\hat{p} \pm z\frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$$

$$.09 \pm 1.96\frac{\sqrt{.09 \cdot .91}}{\sqrt{100}}$$

15-16. Poisson. The distribution of X = number of swirl marks in a polished surface is thought to be Poisson with mean 4.

15. Sketch the z-approximate distribution of X.

DIST OF X $\approx$



$\mu = 4$

POISSON $\Rightarrow \sigma = \sqrt{\mu} = \sqrt{4} = 2$

16. The discrete density of X is $p(x) = e^{-\mu} \frac{\mu^x}{x!}$, x = 0, 1, 2, ... ad inf. Determine the probability of more than 1 swirl mark in a polished surface. Use the discrete probabilities, not the normal approximation and do not give the answer as an infinite sum.

$$P(X>1) = p(1) + p(2) + \cdots + \text{ad inf} = 1 - p(0)$$
$$= 1 - e^{-4}\, 4^0/0! = 1 - e^{-4}$$

17-20. Chi-square. A process has been producing

25% excellent	30% good	40% avg	5% poor

A random sample of 100 chips finds

observed:	18 excellent	24 good	46 avg	12 poor
expected:	25	30	46	12

17. Fill in the expected counts above consistent with past performance.

18. Determine the contribution of category "good" to the chi-square statistic.

$$\frac{(O-E)^2}{E} = \frac{(24-30)^2}{30}$$

19. Determine the degrees of freedom of the chi-square. $k-1 = 4-1 = 3$ "RUBRIC"

NOTE: FROM GEN'L PRINCIPLES # FREE PARAM IN FULL MODEL = 3

FREE PARAM IN H_0 = 0

DIFF = DF = 3

20. Determine p_{SIG} if the chi-square statistic is 13.86.

RIGHT TAIL $df = 3$

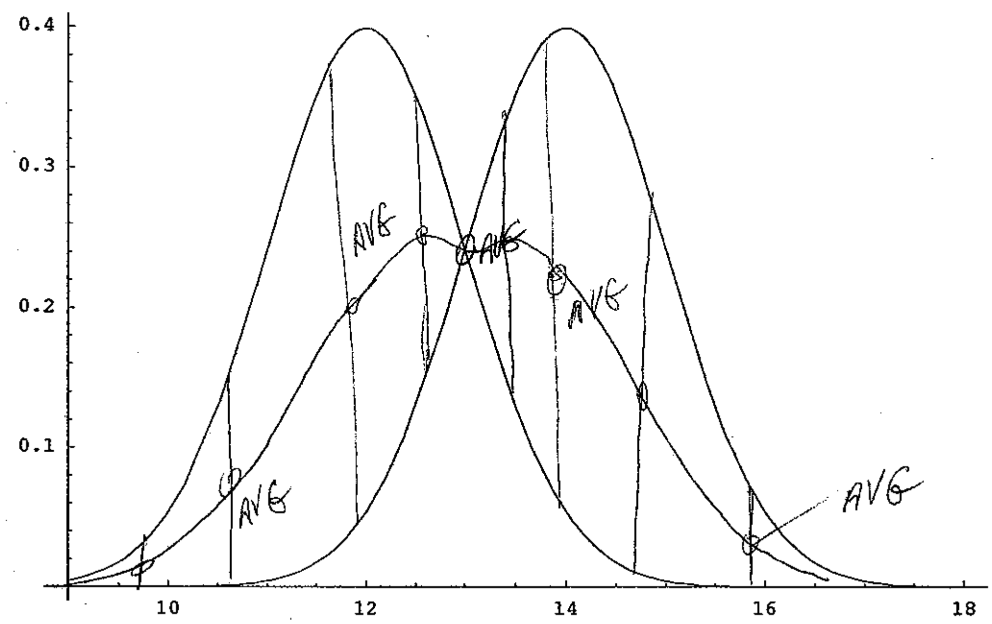
$.005 \leftarrow 12.83$

$\chi^2 = 13.86$

$.001 \leftarrow 16.26$

P_{SIG} IS BETWEEN .001 + .005.

21. Kernel density. Bell curves are placed at each of two points (see below). Plot the kernel density estimate. Take care to do it correctly (show five pts accurately).



22-24. Rules for E, Var, sd. Random variables X, Y are **independent** with

E X = 6 Var X = 4
E Y = 9 Var Y = 2

22. E (2 Y + 4 Y - X + 3) $= 2EY + 4EY - EX + 3$

$= 2(9) + 4(9) - 6 + 3$

23. Var(2 Y + 4 Y - X + 3) $= Var(6Y - X + \cancel{3}) = 36\,Var\,Y + Var\,X$

$= 36(2) + 4$

24. sd (2X - 3Y + $\cancel{4}$) $= \sqrt{Var(2X-3Y)} = \sqrt{4\,Var\,X + 9\,Var\,Y}$

$= \sqrt{4(4) + 9(2)}$

25. Plot regression line. Parts are sampled with-replacement and scored (x, y) where

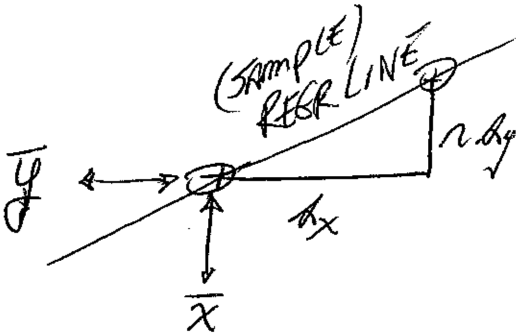
x = serial number of part **y** = hardness.

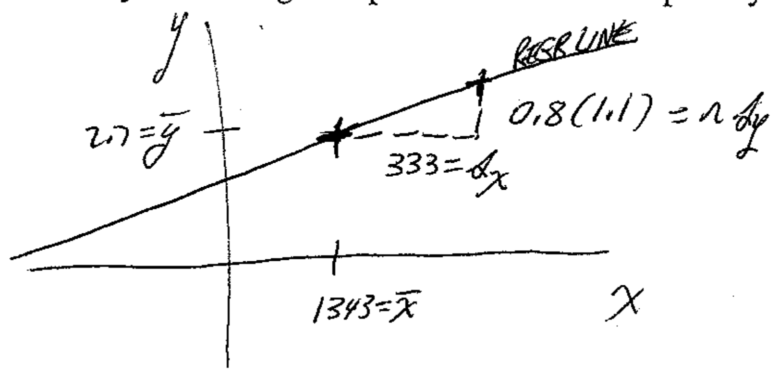
The sample data are:

$\bar{x} = 1343$ $s_x = 333$ **n** = 100 pairs (x, y)

$\bar{y} = 2.7$ $s_y = 1.1$ **r** = 0.8

Sketch a plot the regression line clearly indicating two points on the line explicitly involving the above quantities.





26. Proportionally stratified. A population of motors is stratified by supplier

20% A 10% B 70% C

A **stratified sample** of motors produces the following sample means by stratum

stratum	A	B	C
sample mean	2.4	2.7	2.0

Estimate the population mean μ from the above data.

$$\bar{x} = \sum_{i=1}^{3} w_i \bar{x}_i = .2(2.4) + .1(2.7) + .7(2.0)$$

27. Calculating SD. For the following sample data calculate the sample standard deviation s.

	x	x^2 (i.e. $(x-\bar{x})^2$)
	-3	9
	0	0
	1	1
	2	4
TOT	0	14
AVG	0	

$\bar{x} = 0$

$$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{14}{3}}$$

(OR)
$$s = \sqrt{\frac{n}{n-1}}\sqrt{\overline{x^2} - (\bar{x})^2} = \sqrt{\frac{4}{3}}\sqrt{\frac{14}{4} - 0^2} = \sqrt{\frac{14}{3}}$$

SAME

28-29. Multiple regression. A random sample of 100 of our products is selected from stores nationwide. Each is scored for

y = selling price
x1 = 1 if store is major retailer, 0 if not
x2 = quantity ordered by store

A multiple linear regression is fit to this data resulting in the fitted model

y = 44.75 - 7.80 x1 - 0.83 x2

← x1 (FIX)

28. Determine the regression-based estimate of μ_y if we know from our records
$\mu_{x1} = 0.72$ (i.e. 72% of stores selling our product are major retailers)
$\mu_{x2} = 633$ (i.e. the average quantity ordered per store is 633).

$\hat{\mu}_{y,REGR} = 44.75 - 7.80(0.72) - 0.83(633)$ PREPOSTEROUS NUMBER! WILL ANYONE NOTICE ??

I INTENDED TO USE 0.083 BUT DID NOT CATCH THE ERROR.

29. Compare the estimated margin of error of $\bar{y}$ with that of the estimator (28) if the sample sd of y-scores is 4.55 and the sample multiple correlation is $\hat{R} = 0.77$.

estimated MOE for $\bar{y}$ is $1.96 \frac{s_y}{\sqrt{n}}$

estimated MOE for regression-based estimator (28) is $\sqrt{1-\hat{R}^2}\, 1.96 \frac{s_y}{\sqrt{n}}$

$\sqrt{1-0.77^2}\, 1.96 \frac{4.55}{\sqrt{n}}$

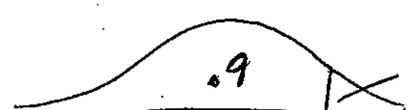
30. t–TEST. A process is in control. Each part produced is score x = finishing time. A sample of 12 will be used to monitor the process in a test of the null hypothesis

H_0: $\mu_x = 5$ (minutes) vs H_1: $\mu_x > 5$ with $\boldsymbol{\alpha = 0.1}$

30. If the test statistic for a sample of 12 evaluates to t = 2.8 what action is taken by the test? Indicate your reasoning.

FOR $\alpha = 0.1$ DF = 11

$t_{CRITICAL} =$



WANT $.1 = \alpha$

$1.363 = t_{CRITICAL}$

TEST REJECTS H_0 FOR LARGE $\bar{x}$

SINCE $t_{STAT} = 2.8 > t_{CRIT} = 1.363$ TABLE IV CUMULATIVE DF $12-1=11$ 90% CUMULATIVE $\boxed{1.363}$

REJECT H_0

(OR) $P_{SIG} = P(t_{11DF} > 2.8) = 0.009$ (TABLE VI) $< \alpha = .10 \Rightarrow$ REJECT H_0