STT 351 TEXT π $\hat{p}$ / OURS p $\hat{p}$ (LIKE MUCH OF "WORLD") 12-6-07

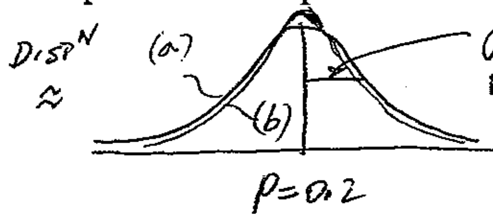
1. A population of resistors has 20% that are below standard.

a. Sketch the approximate distribution of pHAT, the fraction of below standard resistors in a random with-replacement sample of 100 resistors.

$p = 0.2$

$E\hat{p} = p = 0.2$

$SD\ \hat{p} = \dfrac{\sqrt{p(1-p)}}{\sqrt{n}} = \dfrac{0.4}{10} = .04$



$\sigma_{\hat{p}} = \dfrac{\sqrt{p(1-p)}}{\sqrt{n}} = 0.04$

b. Overlay on (a) the corresponding sketch if the population of resistors numbers 500 and the sampling is WITHOUT replacement. $FPC = \sqrt{\dfrac{N-n}{N-1}} = \sqrt{\dfrac{500-100}{500-1}}$

NARROW SD IN (a) TO $\sqrt{400/499}\,(0.04)$

c. Determine the z-approximation of P(pHAT < 0.22) (first find z).

z = STD SCORE OF 0.22 IN DIST^N OF $\hat{p}$ $\dfrac{0.22-p}{\sigma_{\hat{p}}} = \dfrac{0.22-0.2}{.04} = \dfrac{.02}{.04} = \boxed{0.5}$ (z) $\Rightarrow \left[\begin{array}{l} P(Z<.5) \\ = 0.6915 \end{array}\right]$

2. A random with-replacement sample of 40 electronic devices from production finds 13 that are of high grade. Estimate the following:

a. population fraction p of high-grade devices $\hat{p} = \dfrac{13}{40}$

b. population mean mu of scores
x = 1 high grade
x = 0 not

$\mu = E\,X = p\,1 + (1-p)\,0 = p$ ~~$\hat{p}$~~

estimate μ as $\hat{p} = 13/40$

FOR 0-1 SCORES $\mu = p$ (PROB OF 1)

c. sd of pHAT $\sigma_{\hat{p}} = \dfrac{\sqrt{p(1-p)}}{\sqrt{n}}$ EST BY $\dfrac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} = \dfrac{\sqrt{\frac{13}{40}\frac{27}{40}}}{\sqrt{40}}$

---

d. Determine a 90% (not 95%) z-based CI for population p.

TABLE IV

DF	CENTRAL 90%
40	1.684
∞	$\boxed{1.645}$

t DOES NOT APPLY – WE'LL USE $z = 1.645$

BUT TABLE IV SHOWS IT COULD BE OFF FOR $n = 40$

90% z (APPROX) BASED CI $\hat{p} \pm 1.645 \dfrac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} = \dfrac{13}{40} \pm 1.645 \dfrac{\sqrt{\frac{13}{40}\frac{27}{40}}}{\sqrt{40}}$

e. What claim is made for the 90% CI?

$P\left(p \text{ IN } \hat{p} \pm 1.645 \dfrac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}\right) \approx .95$

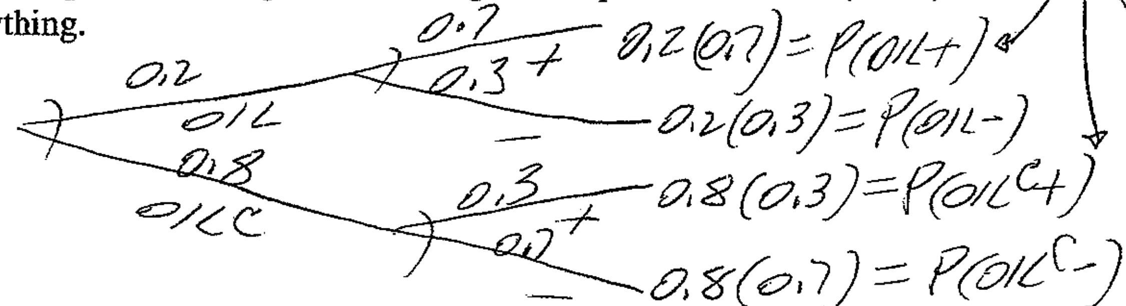
f. Determine the estimated margin of error for pHAT.

EST^D MOE OF $\hat{p}$ IS $1.96 \dfrac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} = 1.96 \dfrac{\sqrt{\frac{13}{40}\frac{27}{40}}}{\sqrt{40}}$

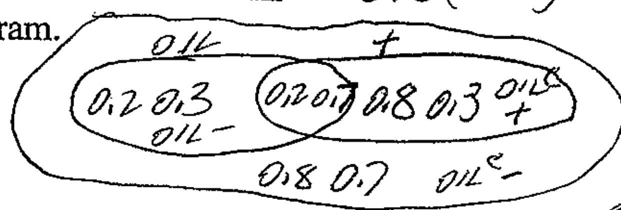
3. P(OIL) = 0.2, P(+ | OIL) = 0.7, P(+ | OIL) = 0.3.

Sym is +

a. Make a complete tree diagram including all endpoints such as P(OIL+). Label everything.



b. Make a complete Venn diagram.



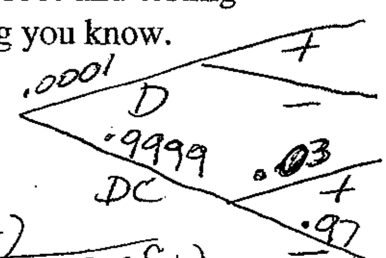
NEEDED HERE

c. Determine P(OIL | +).

$= P(OIL+)/P(+) = (0.2\ 0.7)/(0.2\ 0.7 + 0.8\ 0.3)$

4. One of 10000 persons has a rare disease. A test is available for which the false positive rate is only 3%. A randomly selected person has tested positive. i.e. $P(+|D^c) = .03$

a. Make a tree diagram with disease vs not disease at the root and testing plus or minus on the downstream branches. Fill in everything you know.



b. Determine P(diseased | +) in each of the two cases

P(+ | diseased) = 0 $\Rightarrow P(D|+) = \dfrac{P(D+)}{P(D+)+P(D^c+)}$

$P(D+) = 0\ P(+|D) = 0$

$P(+|D) = P(D+)/P(D) = 0 \qquad = \dfrac{0}{\sim} = 0$

P(+ | diseased) = 1 $\quad P(+|D) = P(D)\,P(+|D) =$

$\Rightarrow P(D|+) = \dfrac{P(D+)}{P(D+)+P(D^c+)} = \dfrac{.0001(1)}{.0001(1) + .9999(.03)} \approx 0$

The answers to (a) and (b) are close.

c. Is the test useful? No − $P(D|+) \sim 0$ EITHER WAY.

(equiv. $P(D^c|+) \sim 1$ WITH GIVEN INFORMATION)

5. A process averages around 4 breakdowns per period. Our process monitor says to stop for inspection and cleaning if we experience more than 7 breakdowns in a given period. Assume that breakdowns follow a Poisson distribution p(x) = e^(-4) 4^x / x! . $e^{-\lambda}\lambda^x/x!\quad x=0,1,2,\ldots$

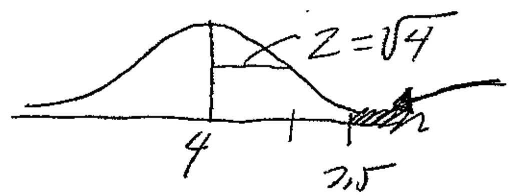
a. Determine P(X > 0) but use 3 for "e" (obviously a bit inexact). Your answer is a number.

$1-p(0) = 1-e^{-4}\frac{4^0}{0!} = 1-e^{-4} \approx 1-\frac{1}{81} = \frac{80}{81}.$

NOT REALLY ACCURATE

b. Because mu = 4 > 3 we have agreed to use a normal approximation for the distribution. Recall that for a Poisson distribution the variance is the mean. Sketch the distribution and shade the area corresponding to x > 7.5 (the region used following the continuity correction when approximating P(X > 7.5)).

P(x) DIST ≈



c. Evaluate z and find the normal approximation to the probability the process will be stopped in a given period.

$z = \frac{7.5-4}{2} = \frac{3.5}{2} = 1.75$ TABLE

ANS $1-.9599 = .0401$

6. 20% of parts are defective. Parts are sampled with-replacement and equal probability. The number of defective parts in a sample n follows the discrete distribution with

p(x) = (binomial coefficient for (n, x)) .2^x (1-.2)^(n-x), x = 0, .., n. $\binom{n}{x}p^x(1-p)^{n-x}$

a. Evaluate p(2) for n = 6.

$\binom{6}{2}.2^2.8^4 = \frac{6!}{2!4!}.2^2.8^4 = \underline{15 .2^2 .8^4}$

b. Evaluate P(X > 2) exactly (not a z-approximation which would be inapplicable here because n = 6 is too small).

$1-p(0)-p(1)-p(2) = 1-\binom{6}{0}.2^0.8^6-\binom{6}{1}.2^1.8^5-\binom{6}{2}.2^2.8^4$
$= 1-.8^6-6.2.8^5-15.2^2.8^4$

c. How many defective parts are expected (on average) in a sample of 6?

$np = 6(.2) = 1.2$ ON AVE

7. Box I contains {R, R, R, G, G, Y} and Box II contains {R, G, G, G, Y}. One box will be chosen, box I with probability 1/3 and box II with probability 2/3. Two balls will be selected WITHOUT replacement from the chosen box.

a. Give P(R1 Y2 | box I is chosen). DRAW W/O REPL FROM RRRGGY

$$P(R1)\,P(Y2|R1) = \frac{3}{6}\,\frac{1}{5}$$

b. Give P(R1 Y2)

$$P(I)\,P(R1\,Y2|I) + P(II)\,P(R1\,Y2|II)$$
$$= \frac{1}{3}\left(\frac{3}{6}\,\frac{1}{5}\right) + \frac{2}{3}\left(\frac{1}{5}\,\frac{1}{4}\right)$$

c. Give P(box I | R1).

$$\frac{P(I\ R1)}{P(I\ R1)+P(II\ R1)} = \frac{1}{3}\left(\frac{3}{6}\right) \Big/ \left(\frac{1}{3}\left(\frac{3}{6}\right)+\frac{2}{3}\left(\frac{1}{5}\right)\right)$$

d. Give P(box I | R1 Y2).

$$\frac{P(I\,(R1\,Y2))}{P(R1\,Y2)} = \frac{\frac{1}{3}\left(\frac{3}{6}\,\frac{1}{5}\right)}{\frac{1}{3}\left(\frac{3}{6}\,\frac{1}{5}\right)+\frac{2}{3}\left(\frac{1}{5}\,\frac{1}{4}\right)}$$

DONE IN (b)

8. 5% of parts are out of spec. Three parts are sampled with replacement.

a. Determine P(out1 in2 in3) .05 .95 .95

P(in1 out2 in3) .95 .05 .95

P(in1 in2 out3) .95 .95 .05

CORRECTIONS

b. From (a) determine P(total of one ~~out~~ OUT and two ~~ins~~ "INS" in three samples).

SUM OF ABOVE

$$= 3(.05)(.95)^2 = \binom{3}{1}.05^1\,.95^2 \text{ (BINOMIAL)}$$

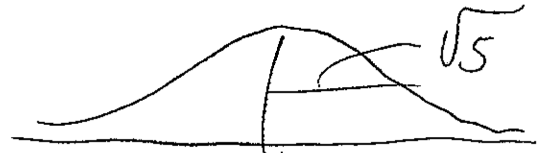
CORRECTION

c. Sketch the approximate distribution of the number X of ~~outs~~ OUTS from a with-replacement sample of 100. Label the mean and sd of the bell curve.

d. Give E X, the expected number of ~~outs~~ OUTS from the sample of 100 in (c). Sketch the Poisson distribution with mean E X. Compare with (c).

NEARLY THE SAME

$n \to \infty$ $p \to 0$



9. Calculate:

a. Sample sd s for data {2, 5, 5}. Do not reduce. $s = \sqrt{\dfrac{(2-4)^2+(5-4)^2+(5-4)^2}{3-1}}$

$\bar{x} = \dfrac{2+5+5}{3} = 4$

b. P(Z in [-1.22, 2.40]) from Z-table.

c. E X^3 for the distribution below.

x	p(x)	$x^3 p(x)$
0	.9	0
2	.1	$8(.1)$

$.8 = E(X^3)$

d. Using your answer to (c) evaluate E ($2(X^3 + 5)$). $= 2(EX^3)+10$ 0.8

10. When a machine begins to squeak it needs lubricating. The times T between such occurrences are modeled as independent and exponentially distributed with a mean of 1.5 hours (P(T > t) = e^(-t / 1.5)) for every t > 0).

a. Determine the probability we wait longer than 3 hours for this to happen. Use e ~ 2.718281828. $P(T>3) = e^{-3/1.5} = e^{-2} = (2.71828\ldots)^{-2}$

b. The process has been running for 5.6 hours without a squeak. Determine the conditional probability that it will run at least 3 more hours without a squeak, **conditional** on the fact that it has already run 5.6 hours without squeaking. You may invoke the "memory less" property of exponential.

$P(T > 5.6 + \underset{\text{ADD'L}}{3} \mid T > 5.6) = P(T>3) = e^{-2}$

c. Determine the probability that a squeak occurs before 2 hours of operation.

$1 - P(T>2) = 1 - e^{-2/1.5}$