Solutions to Practice Questions for Final

Stat 1, Fall 2008

1. $\mu = 430$, $\sigma = 120$ $\qquad X$ = rand. selected students' score

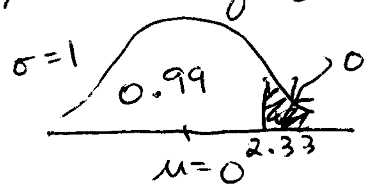
a) $P(X > 600) = P\left(\frac{X-\mu}{\sigma} > \frac{600-430}{120} \right)$

$= P(Z > 1.42)$

$= 1 - P(Z \le 1.42) = 1 - 0.9222$

$= 0.0778$

b) ① top 1% of z-distn



look up 0.99 in body of table, get 2.33

② convert $z = 2.33$ to SAT scale

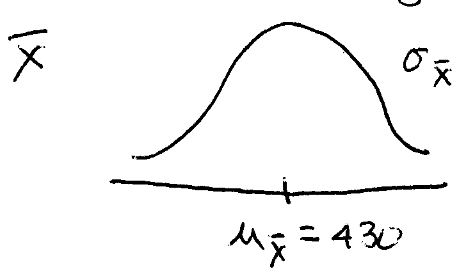
$2.33 = \frac{X-\mu}{\sigma}$

$2.33 = \frac{X-430}{120}$

$430 + 120(2.33) = X$

$X = 709.6 \approx \boxed{710}$

c) $\bar{X}$ will be approx. normally distrib since $n = 36 > 30$. Also, SAT scores are thought to be normally distrib.



$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{120}{\sqrt{36}} = \frac{120}{6} = 20$

$P(\bar{X} > 600) = P\left(\frac{\bar{X}-\mu_{\bar{X}}}{\sigma_{\bar{X}}} > \frac{600-\mu_{\bar{X}}}{\sigma_{\bar{X}}} \right)$

$= P\left(Z > \frac{600-430}{20}\right) = P(Z > 8.5)$

< 0.0002

2. a)

x	\$2	\$4	\$6	−\$6
$p(x)$	1/6	1/6	1/6	3/6

note that $P(X=-6) = P(\text{odd } \#) = 3/6$

b) $\mu = E(X) = \sum_x x \cdot p(x)$

$$= 2(1/6) + 4(1/6) + 6(1/6) + (-6)(3/6)$$

$$= \frac{2+4+6-18}{6} = \frac{-6}{6} = -1$$

$$\sigma^2 = \sum (x-\mu)^2 \cdot p(x)$$

$$= (2-(-1))^2 \cdot (1/6) + (4-(-1))^2 (1/6) + (6-(-1))^2 (1/6) + (-6+1)^2 (3/6)$$

$$= 3^2(1/6) + 5^2(1/6) + 7^2(1/6) + (-5)^2(3/6)$$

$$= 9/6 + 25/6 + 49/6 + 75/6$$

$$= 158/6$$

$$\sigma = \sqrt{158/6} \approx 5.13$$

c) net gain 10,000 games $\approx (-1)(10{,}000)$
$= -10{,}000$ dollars.

This game is a losing prospect,
I wouldn't play it!

3. a) <u>with</u> replacement

B_1 = first bulb burnt out

B_2 = second bulb is burnt out

$$P(B_1 B_2) = P(B_1) \cdot P(B_2 | B_1) = P(B_1) \cdot P(B_2)$$
$$= 5/12 \, (5/12)$$
$$= 25/144 \approx 17\%$$

b) with<u>out</u> replacement

$$P(B_1 B_2) = P(B_1) \cdot P(B_2 | B_1) = 5/\cancel{12}_{3} \cdot (\cancel{4}^{1}/11)$$
$$= \cancel{\frac{20}{132}} = 5/33 = 15\%$$

4. Binomial RV!

X = # of children who liked <u>Finding Nemo</u>

X is binomial $n = 9$, $p = 0.80$

a) $P(X < 5) = P(X \le 4)$ <u>discrete</u> RV

$= 0.020 \leftarrow$ use Table 1

b) $P(5 \le X \le 7) = P(X \le 7) - P(X \le 4)$

$= 0.564 - 0.020$

$= 0.544$

c) X = # who watch Sesame St. out of 10

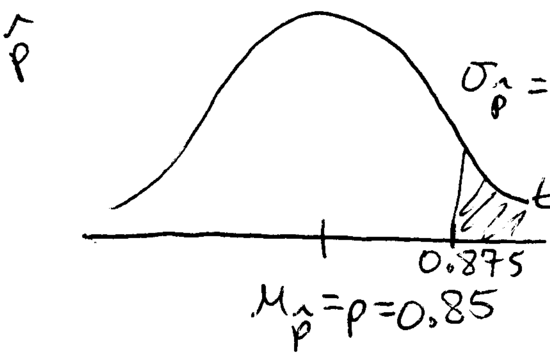
$p = 0.63$ (can't use table 1)

$$P(X = 7) = \binom{n}{k} p^k (1-p)^{n-k} \qquad (k = 7)$$

$$= \binom{10}{7} (0.63)^7 (1-0.63)^{10-7}$$
$$= 120 \, (0.63)^7 (0.37)^3$$
$$= 0.239$$

5. not enough seats will be available if $\hat{p} > \frac{350}{400}$, that is if the sample proportion of booked passengers ~~…~~ is over $350/400$, i.e. more than 350 of the 400 show up.

since $np = 400(0.85) = 340 > 5$
$nq = 400(0.15) = 60 > 5$
} Thus, $\hat{p}$ is approx normal



$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0.85(0.15)}{400}} = 0.01785$$

$$P\left(\hat{p} > \frac{350}{400}\right) = P(\hat{p} > 0.875)$$
$$= P\left(Z > (0.875 - \mu_{\hat{p}})/\sigma_{\hat{p}}\right)$$
$$= P\left(Z > \frac{0.875 - 0.85}{0.01785}\right)$$
$$= P(Z > 1.40)$$
$$= 1 - 0.9192 = 0.0808$$

	M	F	
P	30	20	50
C	45	30	75
J	45	30	75
O	7	8	15
	127	88	215

6. a) $P(\text{Catholic}) = 75/215 \approx 0.349$

b) $P(C|F) = \dfrac{P(C \cap F)}{P(F)}$

$= 30/88 \approx 0.341$

c) $P(C|F) \overset{?}{=} P(C)$

$0.341 \neq 0.349$ (from (a) & (b))

d) $P(J \cap M) = 45/215$ (note unconditional prob.)

e) $P(J \text{ or } M) = P(J) + P(M) - P(J \cap M)$

$= \dfrac{75}{215} + \dfrac{127}{215} - \dfrac{45}{215}$

$= 157/215 \approx 0.73$

f) No! There are subjects who are both Jewish and Male (actually, 45 people).

g) $P(M|J) = \dfrac{P(M \cap J)}{P(J)} = 45/75 = 0.60$

7. a) $\bar{x} = 5$

$s = 3.54$

median $= 4$

b) histogram A

c) no, the Empirical Rule only applies to mound-shaped or normally distrib. data.

d) 1, 3, 100 median $= 3$

mean $= 104/3 \approx 34.7$

e) affected : mean, standard deviation

f) note $\mu \pm 4\sigma$

$75 \pm 4(5) = 75 \mp 20 = (55, 95)$

Chebychev's says at least $1 - \frac{1}{4^2} = 1 - \frac{1}{16} = 93.75\%$ % data within 4 st. dev. of μ.

8. $n = 200$, $x = 120$

a) $\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}$

$\alpha = 1 - 0.92 = 0.08$

$\alpha/2 = 0.08/2 = 0.04$

$z_{0.04} = 1.75$

$\frac{120}{200} \pm 1.75\sqrt{\frac{\frac{120}{200}\left(\frac{80}{200}\right)}{200}}$

$0.60 \pm 1.75\sqrt{\frac{0.6(0.4)}{200}}$

0.60 ± 0.06062

$0.5394 < p < 0.66062$

$0.54 < p < 0.66$

b) Since 0.50 is not in the confidence interval, the coin does not appear to be fair.

9. Type I error = Reject H_0 when H_0 is true = Conclude parachute will open when, in fact, the parachute will NOT open

Type II error = Fail to Reject H_0 when H_a is true = Conclude parachute will not open when, in fact, it will

Possible consequence of Type I error – death

Possible consequence of Type II error – miss out on an exciting jump

~~Worst~~ Type I error is worse so choose α small! $\alpha = 0.01$ and $\beta = 0.05$ is better.

10. a) $$\hat{p}_A - \hat{p}_B \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_A\hat{q}_A}{n_A} + \frac{\hat{p}_B\hat{q}_B}{n_B}}$$

$$\frac{80}{200} - \frac{105}{300} \pm 2.58\sqrt{\frac{\frac{80}{200}\left(\frac{120}{200}\right)}{200} + \frac{\frac{105}{300}\left(\frac{195}{300}\right)}{300}}$$

$$0.40 - 0.35 \pm 2.58\sqrt{0.0012 + 0.000758\bar{3}}$$

$$0.05 \pm 2.58(0.04425)$$

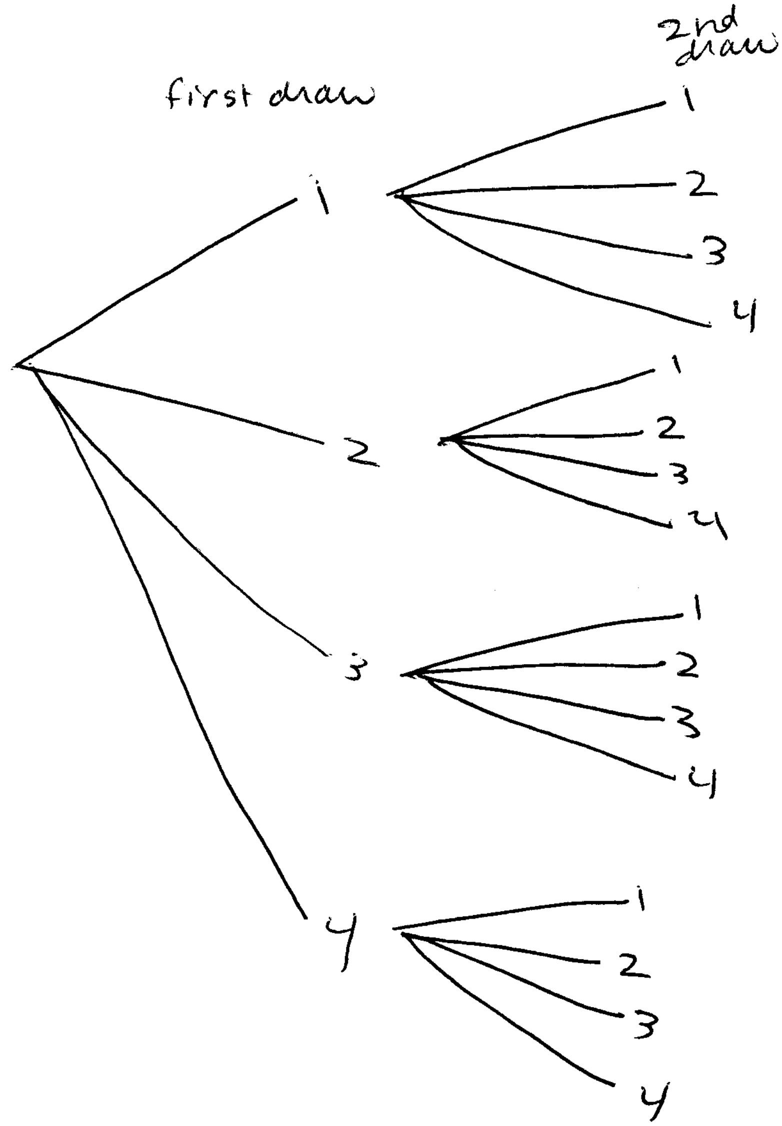
$$0.05 \pm 0.11417$$

$$-0.064 < p_A - p_B < 0.164$$

Since 0 is ~~in~~ in the interval, there does not appear to be a significant difference in the vaccines.

b) No, if we repeatedly drew these size samples & calculated the CI, about 1% of the intervals (in the long run) would fail to contain $p_A - p_B$.

11. a)



{ (1,1) → $\bar{x} = 1$
(1,2) → $\bar{x} = \frac{1+2}{2} = 1.5$
(1,3) → $\bar{x} = \frac{1+3}{2} = 2$
(1,4) → $\bar{x} = \frac{1+4}{2} = 2.5$
(2,1) → $\bar{x} = 1.5$
(2,2) → $\bar{x} = 2$
(2,3) → $\bar{x} = 2.5$
(2,4) → $\bar{x} = 3$
(3,1) → $\bar{x} = 2$
(3,2) → $\bar{x} = 2.5$
(3,3) → $\bar{x} = 3$
(3,4) → $\bar{x} = 3.5$
(4,1) → $\bar{x} = 2.5$
(4,2) → $\bar{x} = 3$
(4,3) → $\bar{x} = 3.5$
(4,4) } → $\bar{x} = 4$

↑
S

b) see tree diagram above

$p(\bar{x} = 1) = 1/16$ one outcome in S that gives $\bar{x} = 1$ out of 16 possible

$\bar{x}$	$p(\bar{x})$
1	1/16
1.5	2/16
2	3/16
2.5	4/16
3	3/16
3.5	2/16
4	1/16
Sum	16/16

c) $\mu = \sum \bar{x} \cdot p(\bar{x}) = 1(1/16) + 1.5(2/16) + 2(3/16) + 2.5(4/16) + 3(3/16) + 3.5(2/16) + 4(1/16)$

$= \frac{1+3+6+10+9+7+4}{16}$

$= \frac{40}{16} = \frac{5}{2} = \boxed{2.50}$

here, $\mu = 2.5 = \mu_{\bar{x}}$

Easier way, $\mu_{\bar{x}}$ = mean of pop drawn from

x	1	2	3	4
p(x)	1/4	1/4	1/4	1/4

12. $n=40$ $\quad \bar{x} = 200$ lb. $\quad s=19$

a) $\bar{x} \pm z_{\alpha/2} \cdot s/\sqrt{n}$

$$200 \pm 1.96\left(19/\sqrt{40}\right)$$

$$200 \pm 1.96(3.00)$$

$$200 \pm 5.89$$

$$200-5.89 < \mu < 200+5.89$$

$$194.11 < \mu < 205.89 \text{ lb.}$$

b) We can be 95% confident the true mean yield of an apple tree is between 194.11 and 205.89 pounds.

c) no, since n $=40$ is "large", $\bar{x}$ will be approx. normally distrib, even if pop. of apple tree yields is not.

d) FALSE!!!

The interval only gives plausible values for μ; it tells us NOTHING about this distribution of the population of apple tree yields we are sampling from.

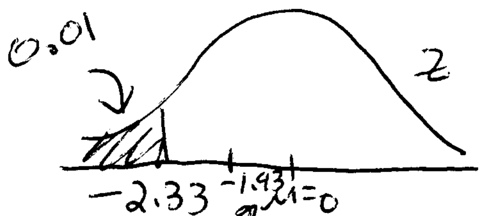
13. $H_0: \mu = 12.3$ min. ($\leftarrow \mu_0$)

$H_A: \mu < 12.3$ min

Test Statistic: $z^* = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}} = \dfrac{12.0 - 12.3}{1.1/\sqrt{50}}$

$= -1.93$

Rejection Region: $\alpha = 0.01$ since $H_a: \mu < 12.3$ left-tail test



Rej. H_0 if $z^* < -2.33$

Since $-1.93 \nless -2.33$, fail to rej. H_0 or "accept H_0".

At the 0.01 level of significance, we cannot conclude the new method reduces the true mean assembly time of the dancing donkey toy.

14. a) St. err. $= s/\sqrt{n}$

$0.0824 = s/\sqrt{62}$

$\sqrt{62}\,(0.0824) = s$

$0.649 = s$

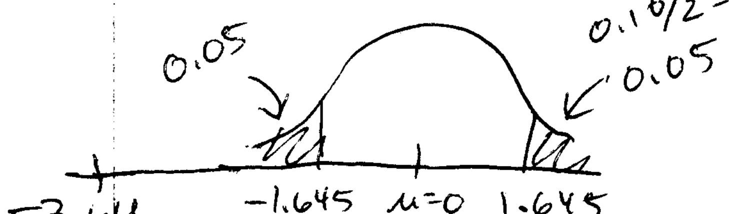
b) $H_0: \mu = 98.6°$

$H_a: \mu \neq 98.6°$

Test stat: $z^* = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} = \dfrac{98.30 - 98.6}{0.649/\sqrt{62}}$

$= -3.64$

Rejection Region: $\alpha = 0.10$, two-tailed (H_a)



Rej. H_0 if $z^* < -1.645$

OR $z^* > 1.645$

$\therefore$ Reject H_0 since

$-3.64 < -1.645$

at the 0.10 l.o.s., we can conclude the ~~true~~ true mean body temp. of healthy adults differs from 98.6° F.

15. ~~a)~~ oops, the letters are a bit off

There is no (a) & (b).

c) -0.80

d) Correlation (in an observational study) does not emply cause & effect. → next page

15d) (continued) Thus, we cannot conclude that having greater access to TV <u>causes</u> longer life expectancy. It is likely that there is a third variable like a measure of the country's wealth or modernization that is causing <u>both</u> longer life & greater access to TV.

e) The points would be more tightly clustered around the line. Since the line has negative slope, it is most likely r changes from -0.80 to -0.86.