SOLUTIONS 60 Points

Statistics 2500, Section 001
Assignment #2: Week of Oct. 5, 2009
Due date: Tuesday, Oct. 20, 2009: 5pm

- The use of Minitab is **required** in the questions indicated. If you think you can use it in other questions, feel free to do so.
- Please **staple** the pages of your assignment together.
- Write your **name**, **lab instructor's name** and **day and time** you attend lab on your assignment.
- Assignments are to be passed into the assignment boxes located just to the left of the math/stats department's general office (HH–3003). Please put your assignment in the box that has your lab instructor's name on it:

  Melissa (Mon. 9am, 10:30am): **BOX 1**
  Vineetha (Mon. 3:30pm, Tues., Thurs. 1:50pm): **BOX 2**
  Chithran (Wed. 9am, 10:30am): **BOX 3**
  Hubert (Fri. 9am, 10:30am): **BOX 4**
  Yunqi (Fri. 2pm): **BOX 5**

- All problem numbers are taken from the textbook *Elementary Statistics* by Bluman and Mayer.

---

1. A study of baseball salaries found that the mean salary of major league baseball players was \$1.7 million, while the median salary of major league baseball players was \$1.2 million.

   (a) What do these results suggest about the shape of a histogram of major league baseball players' salaries?

   (b) If you were the owner of a major league baseball team, which measure of "average" would you emphasize when talking to the public. Why?

   (c) If you were the head of the major league baseball players union (so fighting to get the most for players), which measure of "average" would you emphasize when talking to the public. Why?

2. #18, p. 167.

3. Refer to #34, p. 179 and answer the following:

   (a) #34, p. 179.

   (b) Show whether or not {Being First Year Student} and {Favour Smoking Ban} are independent events.

   (c) Show whether or not {Being Second Year Student} and {Oppose Smoking Ban} are mutually exclusive events.

4. A study revealed that 75% of households have a high-speed internet connection. If a household has an internet connection, 5% use the internet less than 1 hour per week, while 30% use the internet more than 30 hours per week.

   What is the probability that a household has high-speed internet and uses the internet less than 1 hour per week?

5. Refer to #24, p. 212 and do the following.

   (a) #24, p. 212.

   (b) Find the probability of selecting: (i) more than 2 items, (ii) less than 4 items, (iii) at most 6 items.

   (c) Find the mean and standard deviation of the number of items selected.

6. Suppose that 64% of people under the age of 18 own an iPod. A random sample of 7 people under age 18 are selected. Find the following **by hand**:

   (a) What is the probability that 4 of the people own an iPod?

   (b) What is the probability that at most 1 of the people own an iPod?

   (c) What is the probability that at least 2 of the people own an iPod?

   (d) Find the mean number of people under age 18 who own an iPod, and the standard deviation.

7. #51, p. 260.

8. #54, p. 260.

9. #15, p. 268.

10. #25, p. 269.

#1

60 Points

STATS 2500 (001)

SOLUTIONS : ASSIGNMENT #2

(1) (a) Since mean > median, this suggests the histogram would be right-skewed. (1)

(b) If I was the owner, I would use mean as the measure of 'average', because I would like it to appear that players make too much money. (1)

(c) If I led the players' union, I would use the median because I would want it to appear that the owners don't pay the players enough. (1)

---

(2) # 18, p.107.

(a) P(home) = (325 + 406 + 203)/2784 = 934/2784 = 0.34 (1)

(b) ~~P(had or business) = P(had) + P(bus) − P(had and business)~~

~~= 406 + 102 + 732~~

(b) Let A: {ad}, B: {business}.

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

$$= \frac{406+(21)}{2784} + \frac{732+(21+97}{2784} - \frac{1021}{2784}$$

$$= \boxed{\frac{2256}{2784}} = \boxed{0.81}\ (1)$$

(c) A: {1st class}. B: {home}.

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

$$= \frac{325+732}{2784} + \frac{934}{2784} - \frac{325}{2784}$$

$$= \frac{1666}{2784} = \boxed{0.598}\ (1)$$

#3 (a) 34(a). A: {opposes ban}. B: {1st year}.

Word 'given'; conditional probability.

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)} = \frac{27/80}{(15+27+8)/80} = \boxed{\frac{27}{50}} = 0.54\ (1)$$

34(b). A: {favours ban}. B: {2nd year}.

$$P(B|A) = \frac{P(B \text{ and } A)}{P(A)} = \frac{23/80}{(15+23)/80} = \boxed{\frac{23}{38}} = 0.61\ (1)$$

#3 b) A: {1st year}. B: {Favour Bond}.

A, B independent if $P(A \text{ and } B) = P(A) \cdot P(B)$

(no other way for us to establish independence)

$$P(A) \cdot P(B) = \left(\frac{50}{80}\right)\left(\frac{38}{80}\right) = \frac{1900}{6400} = \frac{19}{64}$$ (1)

$$P(A \text{ and } B) = \frac{15}{80} \neq \frac{19}{64}$$

Since $P(A) \cdot P(B) \neq P(A \text{ and } B)$, not independent. (1)

#3 c) A: {1st year}. B: {oppose bond}

A, B mutually exclusive if $P(A \text{ and } B) = 0$.

$$P(A \text{ and } B) = \frac{27}{80} \neq 0.$$ (1)

A, B not mutually exclusive. (1)

#4 A: {high speed}. $P(A) = .75$

If a household → conditional probability.

B: {use less than 1 hr/week}. $P(B|A) = .05$

C: {use more than 30 hrs/wk}. $P(C|A) = .30$

$P(A \text{ and } B) = P(B|A) \cdot P(A)$ (1)

$= (.05)(.75) = \boxed{0.0375}$ (1)

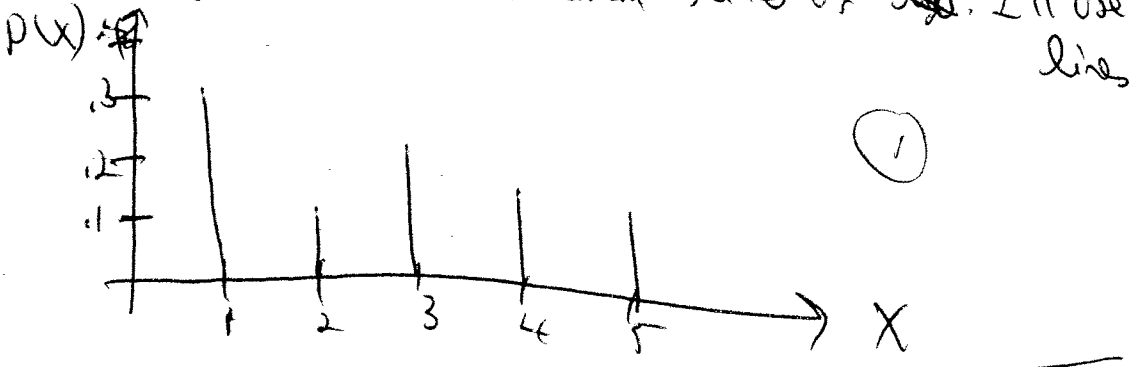
#5 (a) #24, p. 212

Probability distribution: X = number items selected

X	1	2	3	4	5
P(X)	.32	.12	.23	.18	.15

(1)

Graph: Can use vertical lines or bars. I'll use lines.



(1)

#5 b) i) P(more than 2) $= P(X > 2)$

$= P(X=3) + P(X=4) + P(X=5)$

$= 0.23 + 0.18 + 0.15 = \boxed{0.56}$ (1)

(ii) $P(X < 4) = P(X=1) + P(X=2) + P(X=3)$

$= 0.32 + .12 + .23 = \boxed{0.67}$ (1)

(iii) P(at most 6) $= P(X \leq 6)$

$= P(X=0) + \dots + P(X=5) + P(X=6)$

$= \boxed{1}$ (not that $P(X=6) = 0$). (1)

or could do: $P(X \leq 6) = 1 - P(X > 6)$

$= 1 - 0 = \boxed{1}$

c) $\mu = \sum x p(x) = 1(.32) + 2(.12) + 3(.23) + 4(.18) + 5(.15)$

$= \boxed{2.72}$ (2)

$\sum x^2 p(x) = 1^2(.32) + 2^2(.12) + 3^2(.23) + 4^2(.18) + 5^2(.15)$

$= \boxed{9.5}$

$\sigma^2 = \sum x^2 p(x) - \mu^2 = 9.5 - (2.72)^2 = \boxed{2.10}$

$\sigma = \sqrt{2.10} = \boxed{1.45}$ (3)

F

#6 This question involves a binomial random variable. Why?

Not required to state:
(i) fixed number of people ($n=7$)
(ii) 2 outcomes (own iPod or not)
(iii) given $P(\text{own iPod}) = 0.64 = p$
(iv) independent people ('random sample').

X = number that own iPod. $X \sim bin(7, 0.64)$

(a) $P(X=4) = \frac{7!}{(7-4)!4!}(.64)^4(1-.64)^{7-4}$ (1)

$= \frac{7!}{3!4!}(.64)^4(.36)^3$

$= 35(.64)^4(.36)^3 = \boxed{0.274}$ (2)

(b) $P(\text{at most 1}) = P(X \leq 1) = P(X=0) + P(X=1)$ (1)

$= \frac{7!}{0!(7-0)!}(.64)^0(1-.64)^{7-0}$

$+ \frac{7!}{1!(7-1)!}(.64)^1(1-.64)^{7-1}$ (1)

$= (.36)^7 + 7(.64)(.36)^6$

$= .000078 + 0.0098 = \boxed{0.011}$ (2)

(6) c) P(at least 2) = P(X ≥ 2) (1)

= P(X=2) + P(X=3) + -- + P(X=7)

and use binomial formula 6 times.

a very long (but still correct) approach.

Easier: Use complement rule! P(A) = 1 − P(Ā).

P(X ≥ 2) = 1 − P(X < 2)

= 1 − [P(X=0) + P(X=1)]

= 1 − .011 (.011 from (6))

= 0.989 (2)

d) mean = μ = np = 7(.64) = 4.48 (1)

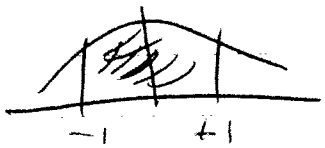
σ = √(np(1−p))

= √(7(.64)(.36))

= √1.61 ≈ 1.27 (1)

(7) #51, p. 260

(i) $P(-1 < Z < 1)$



$= P(Z < 1) - P(Z < -1)$ Using Table E1, E2

$= 0.8413 - .1587 = \boxed{0.6826}$ (1)

Empirical Rule: If distribution bell-shaped, about 68% of observations fall within 1 standard deviation of mean (i.e. in $(\bar{x}-s, \bar{x}+s)$). (1)

The values (.6826 and 68%) are very close.

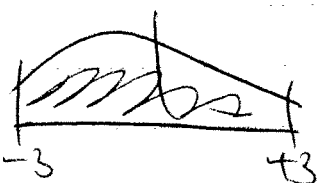
(ii) $P(-2 < Z < 2)$

$= P(Z < 2) - P(Z < -2)$

$= 0.9772 - .0228 = \boxed{0.9544}$ (1)

Again, very close to Empirical Rule (95% of data fall within 2 standard deviations of mean) (1)

(iii) $P(-3 < Z < 3)$



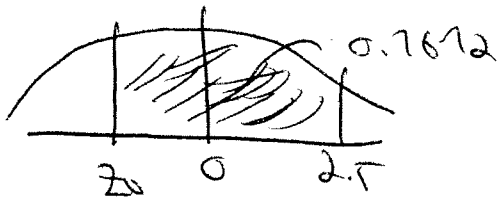
$= P(Z < 3) - P(Z < -3)$

$= 0.9987 - .0013 = \boxed{0.9974}$ (1)

Again, close to 99.7% given in final part of Emp. Rule (1)

(8) #54, p. 260.

Find z_0 such that $P(z_0 < z < 2.5) = 0.7672$.

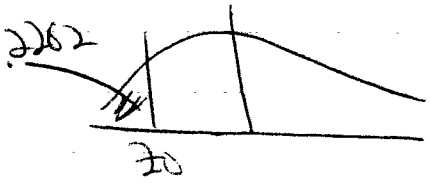


$P(z_0 < z < 2.5) = P(z < 2.5) - P(z < z_0) = 0.7672$

Table E2: $P(z < 2.5) =$ **0.9938** (1)

So: $0.9938 - P(z < z_0) = 0.7672$

$P(z < z_0) = 0.9938 - .7672 =$ **.2266** (1)



Look for 0.2262 in INNER part of E1 or E2.

Then: **$z_0 = -0.75$** (1)

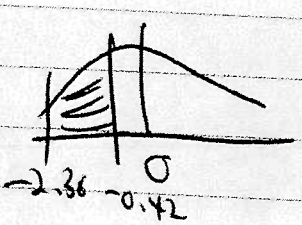
(9) #15, p.268.

X = waiting time. $X \sim N(23.5, 3.6)$

(a) $P(15 < X < 22) = P\left(\frac{15-23.5}{3.6} < Z < \frac{22-23.5}{3.6}\right)$

① for values.

$= P(-2.36 < Z < -0.42)$



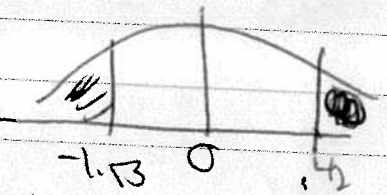
$= P(Z < -.42) - P(Z < -2.36)$

$= 0.3372 - 0.0091 = \boxed{0.3281}$

① for values. ①

(b) $P(X < 18 \text{ or } X > 25) = P(X < 18) + P(X > 25)$

$= P\left(Z < \frac{18-23.5}{3.6}\right) + P\left(Z > \frac{25-23.5}{3.6}\right)$



$= P(Z < -1.53) + P(Z > 0.42)$ ① for values.

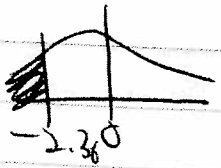
$= 0.063 + [1 - 0.6628] = \boxed{0.4002}$ ①

(0.5) (0.5)

(c) $P(X < 15) = P\left(Z < \frac{15-23.5}{3.6}\right)$

$= P(Z < -2.36)$ ①

$= \boxed{0.0091}$ ①



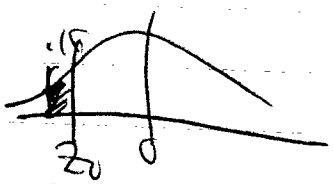
not very likely. ①

(16) #25, p.269

X = length of stay. $X \sim N(5.9, 1.7)$.

(i) 15% of stays less than how many days?

$P(X < x_0) = 0.15$ Find x_0.



Do for z-scores: $P(Z < z_0) = 0.15$

Find 0.15 (or as close as possible) in MIDDLE of table E1 or E2:

$z_0 = -1.04$ (1) (or $z_0 = -1.03$)

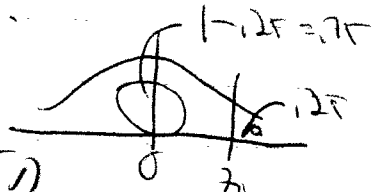
Then: $z = \frac{x - \mu}{\sigma}$

$-1.04 = \frac{x_0 - 5.9}{1.7}$ (1) → $x_0 = -1.04(1.7) + 5.9$

$= 4.13$ days (1)

(ii) 25% Higher (longer) than how many days?

$P(X > x_0) = 0.25$ Find x_0.



z-scores: $P(Z > z_0) = 0.25$

Find .75 in MIDDLE of table: (1) $z_0 = 0.67$ or 0.68

Then: $0.67 = \frac{x_0 - 5.9}{1.7}$ (1) → $x_0 = 1.7(0.67) + 5.9$

(1) $x_0 = 7.38$