STA 621 Midterm – Wednesday, July 15
SHOW ALL WORK

Name: Answer Key

18 pts 1. The following data are the number of home runs for 14 players from the Chicago Cubs in the 1992 baseball season.

$$9, 26, 22, 3, 8, 8, 1, 9, 8, 0, 1, 6, 1, 5$$

$$0, 1, 1, 1, 3, 5, 6, 8, 8, 8, 9, 9, 22, 26$$

$$9 + 26 + \cdots + 1 + 5 = 107$$

$$(9)^2 + (26)^2 + \cdots + (1)^2 + (5)^2 = 1587$$

8 a. Sketch a boxplot for these data.

~~2 pts for sketch~~

~~1 pt for each min + max~~

~~1 pt for each %'ile~~

~~1 for min~~

~~1 for max~~

Min = 0, max = 26

25^{th} %'ile $i = \frac{25}{100}(14) = 3.5$ round up + take 4^{th} value

$\hookrightarrow = 1$

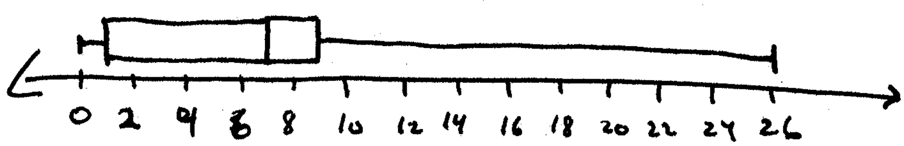
50^{th} %'ile = avg of i^{th} + $(i+1)^{th}$ values

$= \frac{6+8}{2} = 7$ since $i = \frac{50}{100}(14) = 7$

75^{th} %'ile: $i = \frac{75}{100}(14) = 10.5$ round up to 11

take 11^{th} largest

$\hookrightarrow = 9$



4 b. Compute the mode number of home runs by a Cub in 1992.

2 modes (bimodal) 1 and 8

6 b. Based on the 3 standard deviations rule, are there any outliers in this data set? If so, which observations are outliers?

~~2 for mean~~
~~2 for SD~~
~~2 for answer~~

$$s^2 = \frac{\sum x_i^2 - n\bar{x}^2}{n-1} \qquad \bar{x} = \frac{1}{14}(107) = 7.64$$

$$= \frac{1597 - 14(7.64)^2}{13} = 59.17 \quad \Rightarrow s = \sqrt{59.17} = 7.69$$

outlier if z-score is greater than 3 or < -3

check extremes: $0 \rightarrow z = \frac{0 - 7.64}{7.69} = -.99$ not an outlier

$26 \rightarrow z = \frac{26 - 7.64}{7.69} = 2.39$ not an outlier

$\Rightarrow$ No outliers in data set.

15 pts ~~15~~ 13

2. For each of the following variables list all of the adjectives that apply from the following list: Qualitative, Quantitative, Discrete, Continuous.

3 a. Religious affilliation. Qualitative

~~3~~ b. Number of pets owned. Quantitative, Discrete

~~3~~ c. Percentage of body fat. Quantitative, Continuous

~~3~~ d. Eye color. Qualitative

3 e. Shoe size. Quantitative, Discrete

16

3. Two statistics students will be chosen from UGA and 2 will be chosen from Georgia Tech to be the southeastern student delegates at the national statistics convention. There are 200 statistics students at UGA, 88 of whom are women, and there are 300 statistics students at Georgia Tech, 120 of whom are women.

8 a. In how many ways can 2 women delegates be chosen from UGA? from Georgia Tech?

UGA: $\binom{88}{2} = \frac{88!}{86!\,2!} = \frac{88 \cdot 87}{2} = 3828$ ways

GA Tech: $\binom{120}{2} = \frac{120 \cdot 119}{2} = 7140$ ways

8 b. Is it more likely that both of the UGA delegates will be women or that both of the Georgia Tech delegates will be women?

Let G_1 = event that 1st delegate from UGA is woman

G_2 = " " 2nd delegate from UGA " "

T_1 = " " 1st " " Tech " "

T_2 = " " 2nd " " Tech " "

$P(\text{Both UGA delegates female}) = P(G_1 \cap G_2) = P(G_2 | G_1) P(G_1)$

$= \left(\frac{87}{199}\right)\left(\frac{88}{200}\right) = .1924$

$P(\text{Both Tech delegates female}) = P(T_1 \cap T_2) = P(T_2 | T_1) P(T_1)$

$= \left(\frac{119}{299}\right)\left(\frac{120}{300}\right) = .1592$

$\boxed{\text{More likely for UGA}}$

19

4. A company manufactures boxes containing 2 light bulbs, 1 red and 1 green, for the Christmas season. There is a 10% chance that the red light bulb will be defective, and a 20% chance that the green light bulb will be defective. Assume that whether or not the red light is defective is independent of whether or not the green light is defective.

Let X = the number of defective light bulbs in a randomly selected box.

8 a. Find the probability distribution of the random variable X.

X can take values 0, 1 or 2

R = event red defective
G = " green "

$P(X=0) = P(R^c \cap G^c) = P(R^c)P(G^c)$ Because of independence

$= (.9)(.8) = .72$

$P(X=2) = P(R \cap G) = P(R) \cdot P(G) = (.1)(.2) = .02$

$P(X=1) = 1 - P(X=0) - P(X=2) = 1 - .72 - .02 = .26$

Distribution:

x	$f_X(x) = P(X=x)$
0	.72
1	.26
2	.02

5 b. Find the probability that at least one of the light bulbs in a randomly selected box is defective.

$P(X \geq 1) = P(X=1) + P(X=2) = .26 + .02$

$\boxed{= .28}$

6c. Find E(X) and var(X).

$xf(x)$	$x^2 f(x)$
0	0
.26	.26
.04	.08
$E(X) = .3$	.34

$$Var(X) = \sum x^2 f(x) - [E(X)]^2 = .34 - (.3)^2$$
$$= .25$$

16 5. If a test designed to detect when a woman has cancer of type C is applied to a woman who <u>has</u> cancer type C, 95% of the time it will give a positive result, and 5% of the time it will miss the cancer and give a negative result (false negative). If the test is applied to a woman who does <u>not</u> have the cancer, 3% of the time it will give a positive result (false positive), and 97% of the time it will give a negative result. One out of every 100,000 women in the population has cancer type C. If a woman is selected at random from the population,

8 a. What is the probability that the woman tests positive for cancer C?

Let A_1 = event that woman has cancer C

A_2 = event that woman does not have cancer C

B = event that woman tests positive

Given $P(A_1) = \frac{1}{100{,}000}$, $P(B|A_1) = .95$, $P(B|A_2) = .03$

$P(A_2) = 1 - P(A_1) = \frac{99{,}999}{100{,}000}$

$$P(B) = P(B \cap A_1) + P(B \cap A_2) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2)$$
$$= .95\left(\frac{1}{100{,}000}\right) + .03\left(\frac{99{,}999}{100{,}000}\right) = \boxed{.0300092}$$

8 b. If the randomly selected woman tests positive, what is the probability she actually has cancer.

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{P(B|A_1)P(A_1) + P(B|A_2)P(A_2)} = \frac{.95\left(\frac{1}{100{,}000}\right)}{.0300092}$$

from part(a)

$$= \boxed{.0003166}$$

6. Suppose that each morning when my paper delivery boy throws the newspaper toward my driveway, there is a 20% probability that the paper ends up in my thornbush. Suppose that whether or not the paper ends up in the thornbush is independent from one day to the next.

8 a. Find the probability that in a one-week period (7 days), I have to retrieve my paper from the thornbush exactly 3 times.

Binomial exp't $n=7$ $p=.2$

$X = \#$ of successes (papers in thornbush)

$$P(X=3) = \binom{7}{3}(.2)^3(1-.2)^{7-3}$$

$$= \frac{7!}{3!\,4!}(.2)^3(.8)^4$$

$$= 35(.2)^3(.8)^4 = \boxed{.115}$$

8 b. What is the mean for the number of times that the paper is thrown into my thornbush during a one-week period? What is the variance for the number of times the paper lands in the thornbush during a one-week period?

$E(X) = np = 7(.2) = 1.4$

$Var(X) = np(1-p) = 7(.2)(.8) = 1.12$