Quiz II

STAT 305, Section 7
Spring 2020

Instructions

- The quiz is scheduled for 80 minutes, from 02:10 to 03:30 PM. At 03:30 PM the exam will end.
- Total points for the exam is 64. Points for individual questions are given at the beginning of each problem. Show all your calculations clearly to get full credit. Put final answers in the box at the right (except for the diagrams!).
- A forumula sheet is attached to the end of the exam. Feel free to tear it off.
- You may use a calculator during this exam.
- Answer the questions in the space provided. If you run out of room, continue on the back of the page.
- If you have any questions about, or need clarification on the meaning of an item on this exam, please ask your instructor. No other form of external help is permitted attempting to receive help or provide help to others will be considered cheating.
- **Do not cheat on this exam.** Academic integrity demands an honest and fair testing environment. Cheating will not be tolerated and will result in an immediate score of 0 on the exam and an incident report will be submitted to the dean's office.

Name: Key

Student ID: ____________________

Here are some summaries of the data (again using the actual score as the x-value and the person's evaluation of their score as the y-value):

$$\sum_{i=1}^{50} x_i = 1922 \qquad \sum_{i=1}^{50} x_i^2 = 110659$$

$$\sum_{i=1}^{50} y_i = 2954 \qquad \sum_{i=1}^{50} y_i^2 = 179606$$

$$\sum_{i=1}^{50} x_i y_i = 108893$$

(a) Using the summaries above, fit a linear relationship between **the actual score** (x) and **the guessed score** (y).

i. (5 points) Write the equation of the fitted linear relationship.

$\boxed{\hat{y} = b_0 + b_1 X = 63.94 - 0.1266\, X}$

$$b_1 = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sum x_i^2 - n\bar{x}^2}$$

$$= \frac{108893 - 50\left(\frac{1922}{50}\right)\left(\frac{2954}{50}\right)}{\left(110659 - 50\left(\frac{1922}{50}\right)^2\right)} = \frac{-4658.76}{36777.32} = -0.1266$$

$$b_0 = \bar{y} - b_1\bar{x} = \left(\frac{2954}{50}\right) + 0.1266\left(\frac{1922}{50}\right) = 63.95$$

ii. (5 points) Find and interpret the value of R^2 for the fitted linear relationship. $\boxed{R^2 = 11.56\%}$

First find sample correlation (r) & then $R^2 = (r)^2$

$$r = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{\sqrt{\left(\sum x_i^2 - n\bar{x}^2\right)\left(\sum y_i^2 - n\bar{y}^2\right)}} = \frac{108893 - 50\left(\frac{1922}{50}\right)\left(\frac{2954}{50}\right)}{\sqrt{\left[110659 - 50\left(\frac{1922}{50}\right)^2\right]\left[179606 - 50\left(\frac{2954}{50}\right)^2\right]}} = -0.3448$$

$\Rightarrow R^2 = (-0.3448)^2 = 0.1156$

$\Rightarrow$ 11.56% of the variation in the response can be explained by linear relationship between x & y

iii. (5 points) Using the fitted line, what do we suppose a person will guess their score will be if they actually scored a 40.14.

$\boxed{\hat{y} = 58.85}$

$$\hat{y} = 63.94 - 0.1266(40.14)$$

$$= 58.85$$

iv. (2 points) A person who scored a 40.14 on the test predicted that they would score 49.56. What is the residual for this person using the linear relationship?

$e =$ -9.29

$e = y - \hat{y} = 49.56 - 58.85$

$= -9.29$

(b) The JMP output below comes from fitting a quadratic model using the actual score ("`actual_score`") and the square of the actual score (`actual_score^2`).

Response guess_score

Analysis of Variance

	Source	DF	Sum of Squares	Mean Square	F Ratio
SSR	Model	2	4031.5440	2015.77	93.5569
SSE	Error	47	1012.6597	21.55	Prob > F
SST	C. Total	49	5044.2038		<.0001*

Parameter Estimates

	Term	Estimate	Std Error	t Ratio	Prob>|t|
b_0	Intercept	82.626343	1.864461	44.32	<.0001*
b_1	actual_score	-1.246181	0.091868	-13.56	<.0001*
b_2	actual_score^2	0.011013	0.000873	12.62	<.0001*

i. (5 points) Write the equation of the fitted quadratic relationship.

$\hat{y} = 82.6263 - 1.2461(\text{Score}) + 0.011(\text{Score})^2$

ii. (5 points) Find and interpret the value of R^2 for the fitted quadratic relationship.

$R^2 =$ 79.92%

1. $R^2 = \frac{SSR}{SST} = \frac{4031.544}{5044.2038} = 0.7992$

2. $R^2 = \frac{SST - SSE}{SST} = \frac{5044.2038 - 1012.6597}{5044.2038} = 0.7992$

79.92% of the variation in the response variable (the predicted score) can be explained by the quadratic relationship between the response and actual score.

iii. (5 points) Using the fitted quadratic relationship, what do we suppose a person will guess their score will be if they actually scored a 98.74.

$\hat{y} =$ 66.83185

$\hat{y} = 82.6263 - 1.2461(98.74) + 0.011(98.74)^2$

$= 66.83185$

iv. (2 points) A person who scored a 98.74 on the test predicted that they would score 63.55. What is the residual for this person using the quadratic relationship?

$e =$ -3.28

$e = y - \hat{y}$

$= 63.55 - 66.83 = -3.28$

2. Suppose the following is the probability distribution for X.

x	-2	0	1	2	3
f(x)	0.1	a	0.2	0.2	0.3

i. (3 points) Find the value of a that makes this a valid probability distribution. $a =$ 0.2

$\sum f(x) = 1 \implies 0.1 + a + 0.2 + 0.2 + 0.3 = 1$

$\implies a + 0.8 = 1$

$\implies \boxed{a = 0.2}$

ii. (6 points) Calculate the expected value and the standard deviation of X.

E(X)= 1.3

SD(X)= 1.55

$EX = \sum x f(x) = -2(0.1) + 0\overbrace{(0.2)}^{a} + 1(.2)$

$+ 2(0.2) + 3(0.3) = -0.2 + 0 + 0.2 + 0.4 + 0.9$

$= 1.3$

$EX^2 = \sum x^2 f(x) = (-2)^2(0.1) + 0^2(0.2) + 1^2(0.2) + 2^2(0.2) + 3^2(0.3) = 4.1$

$\text{Var } X = EX^2 - [EX]^2 = 4.1 - (1.3)^2 = 2.41$ & $SD(X) = \sqrt{2.41} \approx 1.55$

iii. (2 points) Find the probability that $P(|X| = 2)$

$P(|X| = 2) =$ 0.3

$P(|x| = 2) = P(x = 2 \text{ or } x = -2)$

$= P(x=2) + P(x=-2)$

$= 0.1 + 0.2 = 0.3$

3. As the discussion in the class, King Joffrey was poisoned and Tyrion was found guilty for his death. Tyrion then decided to let the gods decide his fate and demanded a trial by combats. Then it was supposed to have five completely independent combats to decide on his guilt or innnocence. The one who fought for him (his combatant) had $p = 0.7$ probability of success in each combat. Let X be the random variable associated with the number of combats his combatant wins out of the five combats. Let's assume that the probability function for this scenario is defined as

$$P(X = x) = f(x) = \frac{5!}{x!(5-x)!} (0.7)^x \ (0.3)^{5-x} \quad , \quad x = 0, 1, \cdots, 5$$

i. (3 points) What is the probability that his combatant wins **all** combats? $\boxed{p = 0.16807}$

P("combatants wins all combats") $= P(X=5) = f(5)$

$\boxed{\text{Note: } 0! = 1}$

$$= \frac{5!}{5!(5-5)!} (0.7)^5 (0.3)^{5-5} = 0.16807$$

ii. (3 points) What is the probability that his combatant wins **no** combats? $\boxed{p = 0.00243}$

P("combatant wins no combats") $= P(X=0) = f(0)$

$$= \frac{5!}{0!(5-0)!} (0.7)^0 (0.3)^{5-0} = (0.3)^5 = 0.00243$$

iii. (3 points) What is the probability that his combatant wins at least two combats? $\boxed{p = 0.96922}$

P("at least Two wins") $= P(X \geq 2) = f(2) + f(3) + f(4) + f(5)$

Be smart & use compliment!

$$\Rightarrow P(X \geq 2) = 1 - P(X<2) = 1 - P(X \leq 1) = 1 - [P(X=0) + P(X=1)]$$

$$= 1 - [f(0) + f(1)] = 1 - \left[\frac{5!}{0!(5-5)!}(0.7)^0 (0.3)^5 + \frac{5!}{1!(5-1)!}(0.7)^1 (0.3)^{5-1}\right]$$

$$= 1 - [0.00243 + 0.002835]$$

$$= 1 - [0.03078) = 0.96922$$

4. Two discrete random variab1e X and Y can be described using the following probability functions:

x	-1	0	1	2
$f_X(x)$	0.2	0.3	0.3	0.2

y	1	2	3	4
$f_Y(y)$	0.3	0.2	0.2	0.3

i. (4 points) Find the mean and **standard deviation** for X. $\boxed{\mu_X = 0.5 \qquad \sigma_X = 1.025}$

$$\mu_x = E(x) = (-1)(0.2)+(0)(0.3)+(1)(0.3)+(2)(0.2) = 0.5$$

$$E(x^2) = (-1)^2(0.2)+(0)^2(0.3)+(1)^2(0.3)+(2)^2(0.2) = 1.3$$

$$\sigma_x^2 = var(x) = E(x^2) - [E(x)]^2$$

$$= E(x^2) - \mu_x^2$$

$$= 1.3 - 0.5^2 = 1.05$$

$$\sigma_x = \sqrt{1.05} = 1.025$$

ii. (4 points) Find the cumlative distribution function (CDF) of Y.

y	1	2	3	4
$f_Y(y)$	0.3	0.2	0.2	0.3
$F(Y)$	0.3	0.5	0.7	1