EXAM I – STAT 305
Summer 2013

Form A

I have neither given nor received any
unauthorized aid in completing this exam

*Signed*_____

Total score:

General Instructions: Unless otherwise noted you do not need to evaluate quantities to their simplest forms. You can keep answers in a form similar to 19^3 or $\binom{7}{3}$ if that is more convenient for you. You must show work and justify your steps to receive full credit.

1. **True/False**. Circle T if the statement is true and F if the statement is false

   (a) **(T) (F)** There is more than correct way to fit a line through bivariate data.

   (b) **(T) (F)** The standard deviation of a data set is always less than or equal to the mean of the data set.

   (c) **(T) (F)** If data is highly skewed then the sample mean is preferred over the sample median as a measure of "center"

   (d) **(T) (F)** For X, an integer valued random variable, $F_X(2.5) - F_X(1.5) = P(X = 2)$

   (e) **(T) (F)** For Y, an integer valued random variable, $P(Y \geq 3) = 1 - F_Y(3)$

   (f) **(T) (F)** $f(E[X]) = E[f(X)]$ for any random variable X and function f

2. **Short answer**

   (a) Give an example of a random variable that could be modeled using a binomial distribution

   (b) Give an example of a random variable that could be modeled using a poisson distribution

   (c) What is block supposed to accomplish in an engineering experiment?

   (d) Explain why randomization of experimental units to treatments is an important step in experiments.

3. Data was collected on diamonds. There were over 50000 diamonds sampled and multiple variables were collected on this data. We will only consider a sample of 10 of the prices in the data. The following is the sample of diamond prices:
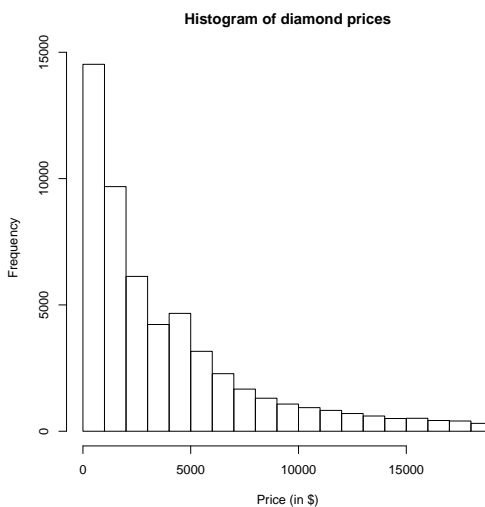
513   570   730   810   906 3672 4054 4197 4241 8037

Some useful things: $\sum_{i=1}^{10} x_i = 27730$ and $\sum_{i=1}^{10}(x_i - \bar{x})^2 = 55815374$

(a) Calculate the sample mean and sample standard deviation.

(b) Calculate the median and interquartile range.

(c) The following is a histogram of the full set of diamond prices

Which is more appropriate to use for your data, the sample mean and variance or the median and IQR? Explain.

4. A waiter noticed that a lot of the tips they received seemed to be such that that the final total was a nice round number. They decided to further examine this by collecting data on the total bill and the tip they received. The following is some of the summaries provided by JMP after fitting a regression predicting TipCents based on BillCents where TipCents = (Tip - floor(Tip)) and BillCents = (Totalbill - floor(Totalbill)).

◢ **Summary of Fit**

RSquare	0.019303
RSquare Adj	0.015251
Root Mean Square Error	0.288629
Mean of Response	0.289262
Observations (or Sum Wgts)	244

◢ **Parameter Estimates**

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	0.3563354	0.035859	9.94	<.0001*
Bill Cents	-0.14385	0.06591	-2.18	0.0300*

(a) What is the fitted least squares line for this data? (Make sure to use the correct variable names - don't just use 'x' and 'y')

(b) What is the sample correlation between BillCents and TipCents? Interpret this value.

(c) What fraction of the raw variability in TipCents is accounted for in the fitting of a line to the data?

(d) Make a prediction for TipCents when BillCents = .27.

(e) What is the mean value for BillCents in the data set?

5. Use the following CDF for X, a discrete random variable, to answer the questions

$$F(x) = \begin{cases} 0 & : x < 0 \\ .4 & : 0 \le x < 1 \\ .6 & : 1 \le x < 2 \\ c & : 2 \le x < 3 \\ 1 & : 3 \le x \end{cases}$$

(a) What must c be if $f(3) = .1$?

(b) What is $P(X > 1)$?

(c) What is $f(1)$?

6. Let X be a discrete random variable that is associated with the number of stars I get on a random song in the game "Guitar Hero".

X	0	3	4	5
f(x)	c	.2	.4	.3

(a) What must c be to make this a valid probability mass function?

(b) What is $P(X \ge 3)$?

(c) What is $E[X]$?

7. Suppose I have a random variable X with probability density function $f(x) = \dfrac{e^{-x}}{1 - e^{-1}}$ for $0 \le x \le 1$ and $f(x) = 0$ for any other value of x.

(a) Find the CDF corresponding to this probability density function.

(b) What is $P(.5 \le X \le 5)$?

(c) What is $P(X = 1/2)$?

(d) What is $E[X]$?

(e) What is $E[e^X]$?

(f) What is $Q(.5)$?