Name: *ANSWER KEY* (108 POINTS) 7/26/2017

Math 173 Introduction to Probability and Statistics Final Exam

1. You are given a *sample* of $n = 5$ measurements:

$$1.2,\ \ 2.5,\ \ 1.7,\ \ 3.1,\ \ 1.5.$$

(a) (2 points) Find the median m.

1.2 1.5 $\boxed{1.7}$ 2.5 3.1

(b) (4 points) Find the mean $\bar{x}$.

$$\bar{x} = \frac{\sum x_i}{n} = \frac{1.2 + 2.5 + 1.7 + 3.1 + 1.5}{5} = \frac{10}{5} = \boxed{2}$$

(c) (4 points) Find the variance s^2 and the standard deviation s.

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1.2	-.8	.64
2.5	.5	.25
1.7	-.3	.09
3.1	1.1	1.21
1.5	-.5	.25

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{.64 + .25 + .09 + 1.21 + .25}{4}$$

$$= \frac{2.44}{4} = \boxed{.61}$$

$$s = \sqrt{s^2} = \sqrt{.61} \approx \boxed{.7810}$$

2. (a) (4 points) An experiment consists of rolling three dice. How many simple events are in the sample space?

$6 \cdot 6 \cdot 6 = \boxed{216}$

(b) (4 points) A businessman in New York is preparing an itinerary for a visit to six major cities. The distance traveled, and hence the cost of the trip, will depend on the order in which he plans his route. How many different itineraries (and trip costs) are possible?

$6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = \boxed{720}$

(c) (4 points) A study is to be conducted in a hospital to determine the attitudes of nurses toward various administrative procedures. If a sample of 5 nurses is to be selected from a total of 25, how many different samples can be selected? (HINT: Is order important in determining the makeup of the sample to be selected for the survey?)

$C^{25}_{5} = \dfrac{25!}{5!\,20!} = \boxed{53{,}130}$

3. (4 points) Your friend has a very interesting movie collection. 70% of her movies are comedies, 50% of her movies star Adam Sandler, and only 20% of her movies are not comedies and do not star Adam Sandler. How many movies does she own that are comedies starring Adam Sandler?

$P(C) = .7$

$P(A) = .5$

$P(C^c \cap A^c) = .2$

$\Rightarrow P(A \cup C) = .8$

$P(A \cup C) = P(A) + P(C) - P(A \cap C)$

$\Rightarrow P(A \cap C) = P(A) + P(C) - P(A \cup C$

$= .5 + .7 - .8$

$= .4$, i.e. $\boxed{40\%}$

4. An experiment can result in events A, B, both A and B, and neither A nor B with the following probabilities.

	B	B^c	TOTAL
A	.34	.46	.8
A^c	.15	.05	.2
TOTAL	.49	.51	1

(a) (4 points) Find $P(A|B)$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{.34}{.49} \approx \boxed{.6939}$$

(b) (4 points) Are events A and B independent? Why/why not?

$\boxed{\text{No,}}$ $P(A|B) \neq P(A)$ ALSO $P(A \cap B) \neq P(A)P(B)$

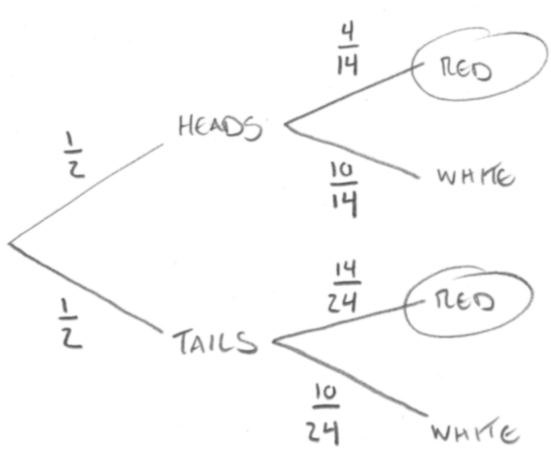
$.6939 \neq .8$ $\qquad .34 \neq (.8)(.49) = .392$

(c) (4 points) Are events A and B mutually exclusive? Why/why not?

$\boxed{\text{No, } P(A \cap B) \neq 0}$

5. Suppose there is a jar that contains 10 white marbles. First, a coin is flipped. If it lands on heads then 4 red marbles are added to the jar. If it lands on tails then 14 red marbles are added to the jar. Second, one marble is selected from the jar.

(a) (4 points) Find the probability of selecting a red marble.



$P(H \cap R) = (\frac{1}{2})(\frac{4}{14}) = \frac{1}{7}$

$P(H \cap W) = (\frac{1}{2})(\frac{10}{14}) = \frac{5}{14}$

$P(T \cap R) = (\frac{1}{2})(\frac{14}{24}) = \frac{7}{24}$

$P(T \cap W) = (\frac{1}{2})(\frac{10}{24}) = \frac{5}{24}$

$$P(R) = \frac{1}{7} + \frac{7}{24} = \frac{24 + 49}{168} = \boxed{\frac{73}{168} \approx .4345}$$

(b) (4 points) Find the probability that the coin landed heads-up, given that a red marble was selected.

$$P(H|R) = \frac{P(H \cap R)}{P(R)} = \frac{\left(\frac{1}{7}\right)}{\left(\frac{1}{7}\right)+\left(\frac{7}{24}\right)} = \frac{24}{24+49}$$

$$= \boxed{\frac{24}{73} \approx .3288}$$

6. City crime records show that 20% of all crimes are violent and 80% are nonviolent, involving theft, forgery, and so on. Ninety percent of violent crimes are reported versus 70% of nonviolent crimes.

(a) (4 points) What is the overall reporting rate for crimes in the city?

$P(V) = .2$

$P(V^c) = .8$

$P(R|V) = .9$

$P(R|V^c) = .7$

LAW OF TOTAL PROBABILITY:

$$P(R) = P(V)P(R|V) + P(V^c)P(R|V^c)$$

$$= (.2)(.9) + (.8)(.7)$$

$$= \boxed{.74 \text{ or } 74\%}$$

(b) (4 points) If a crime in progress is reported to the police, what is the probability that the crime is violent?

$$P(V|R) = \frac{P(V)P(R|V)}{P(R)}$$ (BAYES RULE)

$$= \frac{(.2)(.9)}{(.74)} \approx \boxed{.2432}$$

7. A random variable x has the following probability distribution.

x	0	1	2	3	4
$p(x)$	.08	.12	.34	w	.31

(a) (2 points) Solve for w , i.e. find $p(3)$.

$$\sum p(x) = 1 \implies w = \boxed{.15}$$

(b) (4 points) Find the mean μ, i.e. expected value $E[x]$.

$$\mu = \sum x\,p(x) = 0(.08) + 1(.12) + 2(.34) + 3(.15) + 4(.31)$$

$$= \boxed{2.49}$$

(c) (4 points) Find the variance σ^2 and standard deviation σ.

x	$x-\mu$	$(x-\mu)^2$	$p(x)$
0	-2.49	6.2001	.08
1	-1.49	2.2201	.12
2	$-.49$	.2401	.34
3	.51	.2601	.15
4	1.51	2.2801	.31

$$\sigma^2 = \sum (x-\mu)^2 p(x) = (6.2001)(.08) + (2.2201)(.12) + (.2401)(.34) + (.2601)(.15) + (2.2801)(.31)$$

$$= \boxed{1.5899}$$

$$\sigma = \sqrt{1.5899} \approx \boxed{1.2609}$$

8. Suppose a fair die with faces labeled 1-6 is rolled n times. Let x be the number of times that a 6 is rolled.

(a) (4 points) When $n = 10$, find $P(x \leq 1)$ exactly.

BINOMIAL :

$$P(x=k) = C^n_k \, p^k \, q^{n-k}$$

$$P(x \leq 1) = P(x=0) + P(x=1)$$

$$= C^{10}_0 \left(\tfrac{1}{6}\right)^0 \left(\tfrac{5}{6}\right)^{10} + C^{10}_1 \left(\tfrac{1}{6}\right)^1 \left(\tfrac{5}{6}\right)^9$$

$$= \left(\tfrac{5}{6}\right)^{10} + 10\left(\tfrac{1}{6}\right)\left(\tfrac{5}{6}\right)^9$$

$$\approx \boxed{.4845}$$

(b) (4 points) When $n = 100$, use a normal distribution to approximate $P(x \geq 21)$.

$$\mu = np = 100\left(\tfrac{1}{6}\right) \approx 16.6667$$

$$\sigma = \sqrt{npq} = \sqrt{100\left(\tfrac{1}{6}\right)\left(\tfrac{5}{6}\right)} \approx 3.7268$$

$$P(x \geq 21) \approx P_{NORM}(x \geq 20.5)$$

$$= 1 - P\left(z \leq \frac{20.5 - 16.6667}{3.7268}\right)$$

$$\approx 1 - P(z \leq 1.03)$$

$$\approx 1 - .8485 = \boxed{.1515}$$

9. A random sample of size $n = 49$ is selected from a population with mean $\mu = 53$ and standard deviation $\sigma = 21$.

   (a) (2 points) What will be the approximate shape of the sampling distribution of $\overline{x}$?

   $\boxed{\text{NORMAL}}$ (CENTRAL LIMIT THM)

   (b) (4 points) What will be the mean μ and standard deviation S.E. of the sampling distribution of $\overline{x}$?

   $$\mu = \boxed{53} \qquad S.E. = \frac{\sigma}{\sqrt{n}} = \frac{21}{\sqrt{49}} = \boxed{3}$$

   (c) (4 points) Find the probability that $\overline{x} \geq 57$.

   $$P(\bar{x} \geq 57) = P\left(z \geq \frac{57-53}{3}\right)$$

   $$= 1 - P(z \leq 1.33)$$

   $$= 1 - .9082 = \boxed{.0918}$$

10. (8 points) The first day of baseball comes in late March, ending in October with the World Series. Does fan support grow as the season goes on? Two CNN/USA Today/Gallup polls, one conducted in March and one in November, both involved random samples of 1001 adults aged 18 and older. In the March sample, 45% of the adults claimed to be fans of professional baseball, while 51% of the adults in the November sample claimed to be fans. Construct a 99% confidence interval for the difference in the proportion of adults who claim to be fans in March versus November.

$\hat{p}_1 = .45$

$\hat{p}_2 = .51$

$$S.E. \approx \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

$$= \sqrt{\frac{(.45)(.55)}{1001} + \frac{(.51)(.49)}{1001}} \approx .0223$$

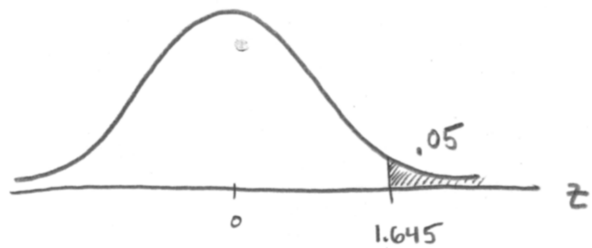
$$p_1 - p_2 \approx \hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2} \, S.E. = .45 - .51 \pm 2.58\,(.0223)$$

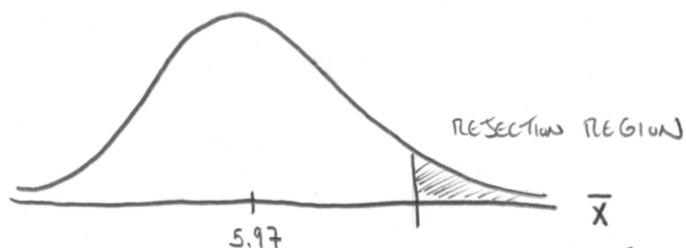
$$= \boxed{[-.1175, -.0025]}$$

8 PNTS. 11. Some sports that involve a significant amount of running, jumping, or hopping put participants at risk for Achilles tendinopathy (AT), an inflammation and thickening of the Achilles tendon. A study in The American Journal of Sports Medicine looked at the diameter (in mm) of the affected tendons for patients who participated in these types of sports activities. Suppose that the Achilles tendon diameters in the general population have a mean of 5.97 millimeters (mm). When the diameters of the affected tendon were measured for a random sample of 31 patients, the average diameter was 9.80 with a standard deviation of 1.95 mm. Is there sufficient evidence to indicate that the average diameter of the tendon for patients with AT is greater than 5.97 mm? Test at the 5% level of significance.

$H_0 : \mu = 5.97$

$H_a : \mu > 5.97$





$X = \mu + 1.645$ S.E.

$= 5.97 + 1.645(.3502)$

$= \underline{\underline{6.5461}}$

$\text{S.E.} \approx \dfrac{s}{\sqrt{n}} = \dfrac{1.95}{\sqrt{31}} \approx .3502$

SINCE $\bar{X} = 9.80 > 6.5461$ ($\bar{X}$ IS IN THE REJECTION REGION)

WE REJECT H_0 IN FAVOR OF H_a.

YES, THERE IS SUFFICIENT EVIDENCE TO CONCLUDE THE AVERAGE DIAMETER OF THE TENDONS FOR PATIENTS WITH AT IS GREATER THAN 5.97 mm

12. Suppose you own an ice cream truck and on 10 consecutive days you record the temperature at noon in °F (x) and the total sales for the day in dollars (y). You obtain the following data.

x	81	83	88	86	92	95	96	88	90	78
y	255	273	312	313	383	411	399	320	318	240

$$\overline{x} = 87.7, \quad \overline{y} = 322.4$$

$$s_x = 5.8699, \quad s_y = 59.0521, \quad s_{xy} = 336.4737$$

4 PNTS. (a) Find the correlation coefficient r. What does it tell you about the relationship between x and y?

$$r = \frac{s_{xy}}{s_x s_y} = \frac{336.4737}{(5.8699)(59.0521)} = \boxed{.9707}$$

STRONG POSITIVE LINEAR RELATIONSHIP.

4 PNTS. (b) Find the equation $y = a + bx$ for the line that best fits the data (i.e. the least-squares line).

$$b = r\left(\frac{s_y}{s_x}\right) = (.9707)\left(\frac{59.0521}{5.8699}\right) = 9.7654$$

$$a = \overline{y} - b\overline{x} = 322.4 - 9.7654(87.7) = -534.0256$$

$$\therefore \quad \boxed{y = -534.0256 + 9.7654\,x}$$

2 PNTS. (c) Use your answer to part (b) to predict what the total sales will be on a day when the temperature at noon is 100°F.

$$y = -534.0256 + 9.7654(100) = \boxed{442.51 \text{ DOLLARS}}$$