10	2
9	10
8	13
7	10
6	5
5↓	7

$\bar{x} = 78$

$M = 81$

ST 305: Exam ~~1~~ 2

By handing in this completed exam, I state that I have neither given nor received assistance from another person during the exam period. I have used no resources other than the exam itself and the basic mathematical functions of a calculator (ie, no notes, electronic communication, notes stored in calculator memory, etc.) Using your calculator for values from probability distributions like the normal or t is OK; however, if you are doing a calculation from a normal distribution show your work all the way to the point of calculating z-scores. I have not copied from another person's paper. I understand that the penalty if I am found guilty of any such cheating will include failure of the course and a report to the NCSU Office of Student Conduct. **I understand that I must show all work/calculations, even if they seem trivial, to get credit for my answers.**

Name: KEY

ID#: ____________________

$$\bar{x} = \frac{1}{n}\sum x_i$$

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

$$Z = \frac{X - \mu}{\sigma}$$

$$r = \frac{\sum \left(\frac{x_i - \bar{x}}{s_x} \right)\left(\frac{y_i - \bar{y}}{s_y} \right)}{n-1}$$

$$b_1 = r\frac{s_y}{s_x}$$

$$b_0 = \bar{y} - b_1\bar{x}$$

$$\text{residual} = y - \hat{y}$$

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

$$P(A^C) = 1 - P(A)$$

$$P(A \text{ and } B) = P(A) \times P(B)$$

$$\mu_X = \sum x_i p_i$$

$$\mu_{a+bX} = a + b\mu_X$$

$$\mu_{X+Y} = \mu_X + \mu_Y$$

$$\sigma_X^2 = \sum \left(x_i - \mu_X \right)^2 p_i$$

$$\sigma_{a+bX}^2 = b^2\sigma_X^2$$

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$$

$$\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2$$

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y$$

$$\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2 - 2\rho\sigma_X\sigma_Y$$

$$\mu_X = np$$

$$\sigma_X = \sqrt{np(1-p)}$$

$$\hat{p} = X / n$$

$$\mu_{\hat{p}} = p$$

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

$$P(X = k) = \frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}$$

$$\mu_{\bar{X}} = \mu$$

$$\sigma_{\bar{X}} = \sigma/\sqrt{n}$$

Definitions. (5 points each) Clearly define each of the following terms.

1. Random variable:

a numerical result from a random phenomenon

2. Law of Large Numbers:

As the sample size n gets large, the sample mean ($\bar{x}$) converges to the pop'n mean(μ)

3. Simple Random Sample:

A sample chosen in such a way that every set of size n has an equal chance of selection.

Multiple Choice. (3 points each) Select the one best answer.

4. Undercoverage occurs when

a. a portion of the individuals in a sample refuse to provide information

B b. a portion of the population is not included in the sampling plan

c. a portion of the individuals in a sample give incorrect or misleading information

5. The number of passengers waiting at a specific bus stop on Monday morning would be an example of

a. a parameter

C b. a lurking variable

c. a discrete random variable

6. If two events A and B are disjoint, then

a. A and B are definitely independent

B b. A and B are definitely not independent

c. A and B may or may not be independent; it depends on the setting

7. Least squares regression should not be used if

a. there is a weak relationship between the explanatory and response variables

B b. there is a nonlinear relationship present in the data

c. the response variable has a skewed distribution

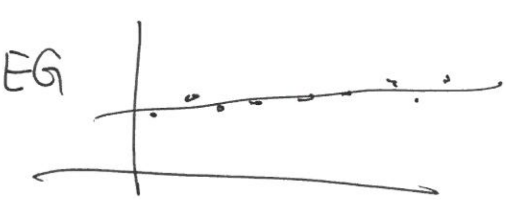
8. A scatterplot with a very flat slope will have

a. a value of r near 1 or -1

B b. a value of b_1 near zero for the least squares line

c. a value of r near zero

EG



For the remaining problems, SHOW YOUR WORK. Numerical answers with no supporting work or explanation will receive zero credit, even if the calculations are trivial.

9. In each of the following experimental designs, one of the three key design principles is absent. Name the missing design element, and briefly explain how you could repair the design.

    a) To learn if a vegetarian diet leads to lower cholesterol levels, a SRS of n=100 vegetarians is selected, and each person's cholesterol is measured. (5 points)

    Control.

    Include a set of individuals eating a non-vegetarian diet (or run a before-after type study where individuals change diets)

    b) To learn if a new inexpensive treatment for kidney disease is as effective as a very expensive kidney transplant, a SRS of 50 patients is selected. One is randomly chosen to get a kidney transplant, while the remaining 49 get the new treatment. At the end of the experiment kidney function is measured on each individual and compared between the transplant and the new treatment. (5 points)

    Replication

    We need multiple individuals to receive Kidney transplants

10. What is an institutional review board (IRB)? (5 points)

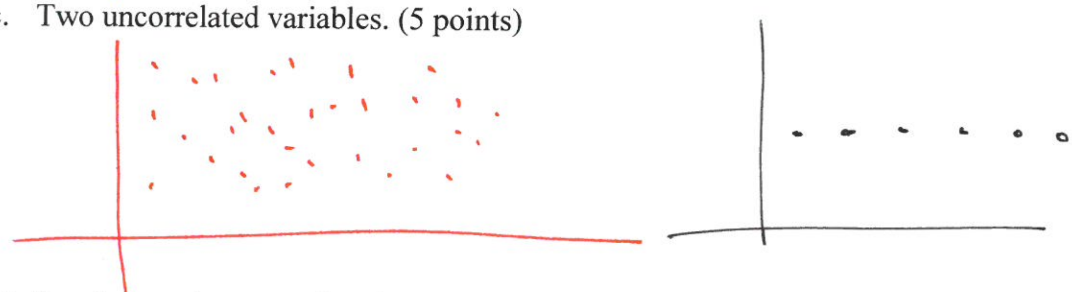
An institutional committee that evaluates the ethical issue surrounding proposed experiments that use human subjects or vertebrate animals.

11. For each of the following, draw a scatterplot consistent with the description.

a. A weak, positive correlation. (5 points)

b. A strong, nonlinear relationship. (5 points)

c. Two uncorrelated variables. (5 points)



12. Rare happenings are often described using the Poisson(λ) distribution, which is a discrete distribution with probabilities defined as $P(X=k)=\dfrac{e^{-\lambda}\lambda^{k}}{k!}$, for positive integer values k = 0, 1, 2, … . Suppose the number of million dollar North Carolina lottery winners in one month follows a Poisson distribution with $\lambda = 2$. Use the indicator variable approach to find the expected number of months with no million dollar winners in a full year. (5 points)

Let T = # of months w/ no winners in 1 year

Define the indicator variables M_i, $i=1,2,\dots 12$ as follows:

$M_i = \begin{cases} 1 & \text{month } i \text{ had no winner} \\ 0 & \text{" " " 1 or more winners.} \end{cases}$ $\boxed{\text{Note: } T = \sum_{i=1}^{12} M_i}$

Using the formula for the Poisson distn,

$P(M_i = 1) = \dfrac{e^{-2}2^{0}}{0!} = e^{-2} \approx .135$

$E(M_i) = 0(1-.135) + 1(.135)$
$= .135$

$\Rightarrow E(T) = E\{\sum M_i\}$
$= \sum\{E(M_i)\}$
$= 12(.135) \doteq \boxed{1.62}$ *

13. The total cost for a semester at a local college is \$200 tuition per credit hour taken, plus \$1,000 in required fees (no matter how many credits are taken). The number of credit hours taken by students at the college has mean 15 and standard deviation 3.

a. Find the mean and standard deviation of the total cost (tuition plus fees) per semester at the college. (5 points)

$\ast$ $\boxed{= 4{,}000}$

Let X = # credits taken. $\Rightarrow$ Total $= 1{,}000 + 200X$

$E(x) = E(1{,}000 + 200x) = 1000 + 200\,E(x) = 1000 + 200(15)$

$V(x) = (200)^2 V(x) = 40{,}000(3^2) = 360{,}000$

$\sigma_x = \sqrt{360{,}000} = \boxed{600}$ $\ast$

b. A sorority house has 36 women living in it. Assume that they can be considered a SRS of students from the college. What is the probability that their average total costs in one semester is less than \$3,800? (5 points)

If X_i = cost for woman i, then $\bar{X}$ = avg cost for all 36

Need $P(\bar{X} < 3800)$

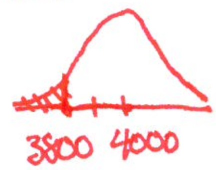
- $\mu_{\bar{X}} = \mu_X = 4000$ <from part a>
- $\sigma_{\bar{X}} = \sigma_X/\sqrt{36} = \frac{600}{6} = 100$
- since $n = 36$ is "largish", use CLT

$P(\bar{X} < 3800)$

$= P\left(Z < \frac{3800 - 4000}{100}\right)$

$= P(Z < -2)$

$= \boxed{.0228}$ $\ast$



14.

a. What are the requirements to have a binomial setting? (5 points)

i) sample size, n, is not random
ii) outcomes of n trials are indep
iii) $P(\text{success}) = p$ for all trials
iv) Each outcome is either success or failure

b. In a binomial setting, what random variable has a binomial distribution? (5 points)

X = count of successes

15. A professional golfer scores a birdie on 10% of her holes (assume the scores for each hole are independent)

a. Find the probability that she makes exactly 5 birdies in a standard (18-hole) round of golf. (5 points) Let X = # of birdies in 18 holes

$X \sim B(18, .10)$

$P(X=5) = \binom{18}{5}(.10)^5(.90)^{13} = \boxed{.0218}$

b. Find the mean and standard deviation of the number of birdies she scores per round. (5 points)

$\mu_x = np = (18)(.10) = 1.8$

$\sigma_x = \sqrt{np(1-p)} = \sqrt{10(.1)(.9)} = \sqrt{1.62} = 1.27$

c. If she plays 200 rounds of golf in a year, find the probability that she averages more than 2 birdies per round. (5 points)

Need: $P(\bar{X} > 2)$, where $\bar{X}$ = avg # of birdies in n = 200 rounds.

Since n = 200 is large, we use CLT:

$\bar{X} \,\dot{\sim}\, N(\mu_{\bar{X}}, \sigma_{\bar{X}})$, where $\mu_{\bar{X}} = \mu_X = 1.8$

$\sigma_{\bar{X}} = \sigma_X/\sqrt{n} = 1.27/\sqrt{200} = .09$

$\Rightarrow \bar{X} \,\dot{\sim}\, N(1.8, .09)$

$P(\bar{X} > 2) = P\left(Z > \frac{2-1.8}{.09}\right) = \cancel{2.22}\; P(Z > 2.22)$

$= \boxed{.0132}$ **