Exam II

STAT 105, Section A
SPRING 2016

Instructions

- The exam is scheduled for 80 minutes, from 8:00 to 9:20 AM. At 9:20 AM the exam will end.
- 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: ______________________________

Student ID: ______________________________

1. An engineer for City Bikes, a rent-and-return bike company, is working on decreasing the time it takes to bring a bicycle to a complete halt. The company is interested in fitting bikes with brake that is has consistently short brake times. Bike mechanisms are durable and rarely wear out, but the rubber brake pads often do and may be replaced locally (i.e., the type of brake pad used is out of the company's control, but the brand of brake mechanism is within their control). His goal is to recommend a brake mechanism that has low stopping times regardless of the brake pads used. He is looking at three different brands of brake mechanism and three common types of rubber brake pads and has decided to use a factorial study to determine which brake mechanism is best. The engineer assigns the brake mechanism brand to Factor A and the rubber pads to Factor B. Stopping speeds for each combination of brake mechanism and brake pads were tested four times under similar condtions.

   The results are recorded below.

	Brake Pads		
Brand	Type 1	Type 2	Type 3
Brand 1	8.73	8.47	9.75
	9.16	5.72	10.93
	9.54	7.36	11.56
	8.02	5.72	12
Brand 2	5.36	4.53	15.01
	5.39	4.29	16.04
	5.16	4.01	14.74
	5.26	5.85	16.24
Brand 3	9.2	7.52	9.15
	8.55	6.74	10.97
	8.14	6.59	10.19
	7.49	5.22	11.98

   The following summaries may help in this problem:

	Brake Pad			
Brand	Type 1	Type 2	Type 3	
Brand 1		$\bar{y}_{12} = 6.82$	$\bar{y}_{13} = 11.06$	
Brand 2	$\bar{y}_{21} = 5.29$		$\bar{y}_{23} = 15.51$	$\bar{y}_{2\cdot} = 8.49$
Brand 3	$\bar{y}_{31} = 8.35$	$\bar{y}_{32} = 6.52$	$\bar{y}_{33} = 10.57$	$\bar{y}_{3\cdot} = 8.48$
	$\bar{y}_{\cdot 1} = 7.5$		$\bar{y}_{\cdot 3} = 12.38$	$\bar{y}_{\cdot\cdot} = 8.63$

   (a) (2 points) Report the value of $\bar{y}_{1\cdot}$

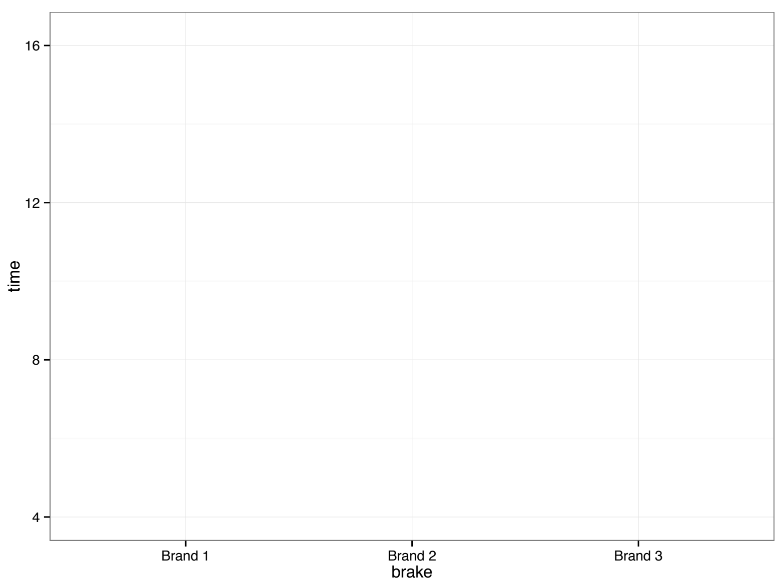
   (b) (2 points) Report the value of $\bar{y}_{\cdot 2}$

(c) (3 points) Find the fitted main effect of braking using each brand, a_1, a_2, and a_3, that you would get from factorial model that ignores interactions.

(d) (3 points) Ignoring possible interactions, give the estimated values $\hat{y}_{22}$ and $\hat{y}_{23}$.

(e) (2 points) How do the estimated values computed above compare to the average for the same combinations seen in the data? Does it appear that ignoring interactions was a good choice?

(f) (5 points) Using the template below, create a profile plot for this data:



(g) (2 points) Using the plot does it appear that there is an interaction between brake mechanism and rubber type? If you had to recommend a brake mechanism and had to consider low stopping times and consistency across brake pads which would you suggest?

2. Two brands (A and B) of electrical circuit boards are being used by a company. The circuits each have five connectors which can either be defective or non-defective (meaning that each circuit board could have 0, 1, 2, 3, 4, or 5 nondefective connectors). The company knows that the brand matters when it comes to the rate with which the connectors are defective. Unfortunately, all the circuit boards have been placed in the same container and there is no clear way to tell them apart.

   We do know the following though:

$$P(\text{“There are } k \text{ defective connectors on the board”} \,|\, \text{“The board is brand A”}) = \frac{5!}{(5-k)!k!}\left(\frac{1}{4}\right)^k\left(\frac{3}{4}\right)^{5-k}$$

$$P(\text{“There are } k \text{ defective connectors on the board”} \,|\, \text{“The board is brand B”}) = \frac{5!}{(5-k)!k!}\left(\frac{3}{4}\right)^k\left(\frac{1}{4}\right)^{5-k}$$

   Suppose that we also know that there were 300 boards of brand A and 200 boards of brand B.

   In this problem, $f(x, w) = P(X = x, Y = y)$ defines the joint probability function.

   (a) (2 points) Find the probability that a randomly chosen board is from brand A.

   (b) (2 points) Find the probability that a randomly chosen board has no defective connectors *given* that it is from brand A.

   (c) (2 points) Find the probability that a randomly chosen board is from brand A and has 0 defective connectors.

   (d) (2 points) Find the probability that a randomly chosen board is from brand B and has 0 defective connectors.

   (e) (2 points) Find the probability that a randomly chosen board has 0 defective connectors.

(f) (2 points) Find the probability that a randomly chosen board is from brand A *given* that it has 0 defective connectors.

(g) (2 points) Find the probability that a randomly chosen board is from brand B *given* that it has 2 defective connectors.

3. (a) (2 points) Let X be a binomial random variable with $n = 6$ and $p = 0.4$. Find $P(X = 3)$

   (b) (2 points) Find $E(X)$

   (c) (2 points) Find $Var(X)$

4. (a) (2 points) Let Y be a binomial random variable with $n = 20$ and $E(Y) = .1$. Find the value of p for this distribution.

   (b) (2 points) Let W be a poisson distribution with $P(W = 0) = .2$. Find λ for this distribution.

5. (4 points) Suppose that $U \sim poisson(\lambda_1)$ and $V \sim poisson(\lambda_2)$. If $P(U = 1) = P(V = 1)$ does that mean that $\lambda_1 = \lambda_2$? Justify your answer using any relevant equations (hint: though not required, it may help you to plot the probability equation different values of λ).

6. After winning an enormous sum on a (rigged) casino dice game Danny Ocean moves on to another game. He has rigged one of the casino's slot machines so that it will land on "triple cherries" (in which case he wins \$50,000) with probability 0.2 and "Triple 7s" (in which case he will win \$2,000,000) with probability 0.05. Considering either of these outcomes to be a "win" he defines the following random variables in order to do some probability calculations on the back of a cocktail napkin:

   - W: whether he wins or loses a given attempt. If he wins a given attempt then $W = 1$ and if he loses $W = 0$.
   - K: the amount of he wins on a given attempt in dollars (it could be 0, 50,000 or 200,000)
   - T: the number of times he has to play before he gets "Triple 7s" for the first time.

   He plans to stop playing as soon as he gets "Triple 7s" (because if he got it twice it would draw serious suspicion).

   (a) (2 points) Provide the probability function for W.

   (b) (2 points) Find the expected value of W.

   (c) (2 points) Find the variance of W.

   (d) (2 points) Find $\mathbb{E}(K)$.

   (e) (2 points) Find $\mathbb{E}(T)$.

   (f) (2 points) What is the probability that Danny will get "Triple 7s" for the first time on his 2nd attempt?

(g) (4 points) What is the probability that it will take more than 3 games for Danny to win for the first time?

(h) (4 points) Danny gets the "Triple 7s" jackpot, but decides to play three more games - what is the probability that he wins at least one of them?

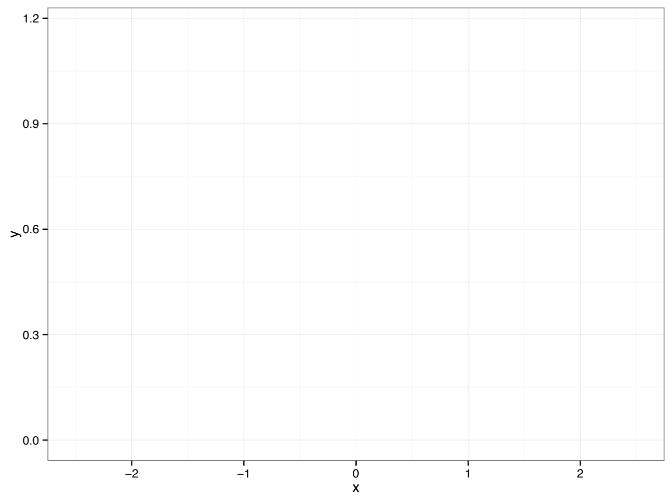
7. Suppose that X is a continuous random variable with probability density function (pdf):

$$f(x) = \begin{cases} 0.5 + cx^2 & -2 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

where c is a constant (not necessarily positive).

(a) (2 points) What is the value of c if $f(x)$ is a valid probability density function?

(b) (5 points) Sketch the probability density function using the grid below (including the points $(-2, f(2))$ and $(2, f(2))$).



(c) (4 points) What is the cumulative density function, $F(x)$?

(d) (2 points) What is the probability that X takes a value greater than 1?

(e) (2 points) What is the probability that X takes a value between 0 and 1?

STAT 105 Exam II
Reference Sheet

'actorial Analysis (Two Factors)

.ssuming

- Factor A with levels $1, 2, ..., I$,
- Factor B with levels $1, 2, ..., J$,
- n is the total number of observations,
- n_{ij} is the total number of observations with Factor A at level i and Factor B at level j,
- $n_{i\cdot}$ is the total number of observations with Factor A at level i,
- $n_{\cdot j}$ is the total number of observations with Factor B at level j.
- y_{ijk} is the kth observation where Factor A is at level i and Factor B is at level j.

$$y_{ij\cdot} = \sum_{k=1}^{n_{ij}} y_{ijk} \qquad \bar{y}_{ij} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} y_{ijk}$$

$$\bar{y}_{i\cdot} = \frac{1}{J} \sum_{j=1}^{J} \bar{y}_{ij} \qquad \bar{y}_{\cdot j} = \frac{1}{I} \sum_{i=1}^{I} \bar{y}_{ij}$$

$$\bar{y}_{\cdot\cdot} = \frac{1}{I} \sum_{i=1}^{I} \bar{y}_{i\cdot} = \frac{1}{J} \sum_{j=1}^{J} \bar{y}_{\cdot j}$$

Main effect of Factor A at level i	$a_i = \bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}$
Main effect of Factor B at level j	$b_j = \bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot}$
Interaction of Factor B at level j and Factor A at level i	$ab_{ij} = \bar{y}_{ij} - a_i - b_j + \bar{y}_{\cdot\cdot}$
Fitted Value (no interactions)	$\hat{y}_{ij} = a_i + b_j + \bar{y}_{\cdot\cdot}$
Fitted Value (including interactions)	$\hat{y}_{ij} = a_i + b_j + ab_{ij} + \bar{y}_{\cdot\cdot}$

3asic Probability Rules

'robability A given B	$P[A\|B] = \frac{P[A,B]}{P[B]}$
'robability A and B	$P[A, B] = P[A\|B]P[B] = P[B\|A]P[A]$
'robability A or B	$P[A \text{ or } B] = P[A] + P[B] - P[A, B]$

)iscrete Random Variables

'robability function	$P[X = x] = f_X(x)$
'umulative probability function	$P[X \leq x] = F_X(x)$
;xpected Value	$\mu = E(X) = \sum_x x f_X(x)$
'ariance	$\sigma^2 = Var(X) = \sum_x (x - \mu)^2 f_X(x)$
tandard Deviation	$\sigma = \sqrt{Var(X)}$

Geometric Random Variables

X is the trial count upon which the first successful outcome is observed performing independer trials with probability of success p.

Possible Values	$x = 1, 2, 3, \ldots$
Probability function	$P[X = x] = f_X(x) = p^x(1 - p)^{x-1}$
Expected Value	$\mu = E(X) = \frac{1}{p}$
Variance	$\sigma^2 = Var(X) = \frac{1-p}{p^2}$

Binomial Random Variables

X is the number of successful outcomes observed in n independent trials with probability of success p.

Possible Values	$x = 0, 1, 2, \ldots, n$
Probability function	$P[X = x] = f_X(x) = \frac{n!}{(n-x)!x!} p^x(1 - p)^{n-x}$
Expected Value	$\mu = E(X) = np$
Variance	$\sigma^2 = Var(X) = np(1 - p)$

Poisson Random Variables

X is the number of times a rare event occurs over a predetermined interval (an area, an amour of time, etc.) where the number of events we expect is λ.

Possible Values	$x = 0, 1, 2, 3, \ldots$
Probability function	$P[X = x] = f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}$
Expected Value	$E(X) = \lambda$
Variance	$Var(X) = \lambda$

Continuous Random Variables

Probability density function	$P[a \leq X \leq b] = \int_a^b f_X(x)dx$
Cumulative probability function	$P[X \leq x] = F_X(x) = \int_{-\infty}^x f_X(t)dt$
Expected Value	$\mu = E(X) = \int_{-\infty}^{\infty} x f_X(x)dx$
Variance	$\sigma^2 = Var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f_X(x)dx$
Standard Deviation	$\sigma = \sqrt{Var(X)}$