

Exam II

STAT 105, Section B FALL 2015

Instructions

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

Brand	Brake Pads		
	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

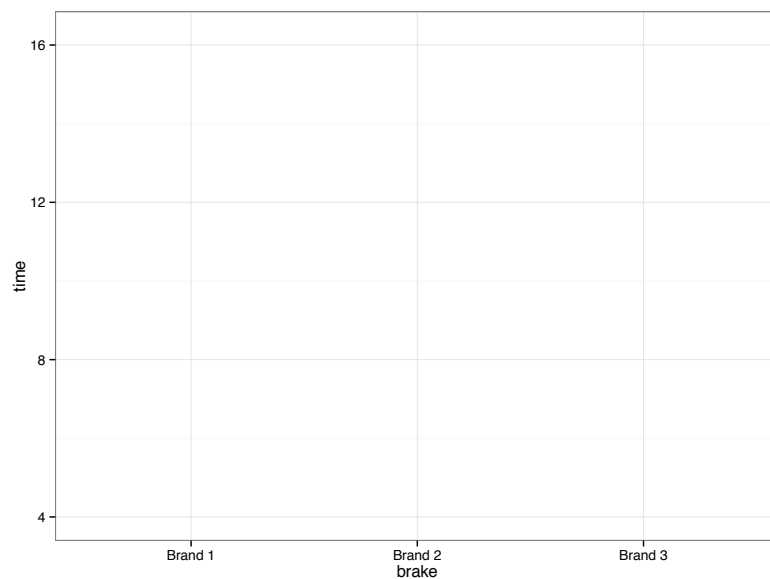
The following summaries may help in this problem:

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

(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. After winning an enormous sum on a (rigged) casino dice game Danny Ocean moves on to another game. In roulette, a ball bounces on a spinning wheel with 38 pockets numbered 0, 00, 1, 2, ..., 36, and bets are made on which pocket the ball will eventually come to rest. In a fair game, the ball has the same chance of coming to a rest in every pocket, but Mr. Ocean has rigged the game (using metallic dust and magnets). In this rigged version, the ball will come to rest in 00 with probability 0.14, 0 with probability 0.14, and any other of the 36 pockets with probability 0.02. Mr. Ocean will win his bet if the ball lands in either 00 or 0.

He defines two random variables: X (which takes the value 1 if he wins the first spin and 0 if he loses), and T , the number of attempts needed to win for the first time.

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

(b) (2 points) Find the variance and expected value of X .

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

(d) (2 points) What is the probability that Danny will win for the first time on his 5th attempt?

(e) (4 points) What is the probability that it will take more than 2 games for Danny to win his first bet?

(f) (4 points) Danny wins, but decides to play five more games - what is the probability that he wins at least one of them?

3. Let X be a normal random variable with a mean of 2 and a variance of 9 (i.e., $X \sim N(2, 9)$) and let Z be a random variable following a standard normal distribution.

(a) Find the following probabilities (note: Table B-3 may be helpful):

i. (2 points) $P(Z \leq 2)$

ii. (2 points) $P(|Z| \geq 1)$

iii. (2 points) $P(0 \leq Z < 3)$

iv. (2 points) $P(X < 3)$

v. (2 points) $P(|X| \leq 4.5)$

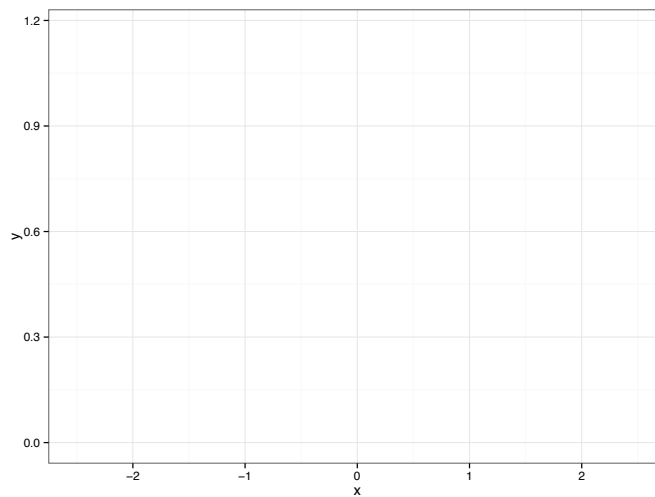
(b) (5 points) Find the value a so that $P(-a + 2 < X < a + 2) = .95$ (approximate as needed).

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

$$f(x) = \begin{cases} 0 & x < 0 \\ cx^2 & -2 < x < 2 \\ 0 & x \geq 2 \end{cases}$$

where c is a constant.

- (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?

5. Suppose we know that W is a binomial random variable with $n = 5$ trials, each with the same probability of success p , but that the value of p is depends on another random variable X . X will take one of three values, $\frac{1}{4}$, $\frac{2}{4}$, or $\frac{3}{4}$, each with the same probability. The value X takes will then serve as the value for the probability of success, p , for W , i.e., if $X = \frac{1}{4}$ then W is a Binomial($5, \frac{1}{4}$). Two of the conditional probability functions that result from this arrangement are below:

$$f(w|X = \frac{1}{4}) = \begin{cases} \frac{5!}{(5-w)!w!} (\frac{1}{4})^w (1 - \frac{1}{4})^{5-w} & w = 0, 1, 2, 3, 4 \text{ or } 5 \\ 0 & \text{otherwise} \end{cases} \quad f(w|X = \frac{2}{4}) = \begin{cases} \frac{5!}{(5-w)!w!} (\frac{2}{4})^w (1 - \frac{2}{4})^{5-w} & w = 0, 1, 2, 3, \\ 0 & \text{otherwise} \end{cases}$$

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

- (a) (2 points) Find the conditional probability function $f(w|X = \frac{3}{4})$.

- (b) (2 points) Find the joint probability $f(\frac{1}{4}, 0)$.

- (c) (2 points) Find the joint probability $f(\frac{2}{4}, 0)$.

- (d) (3 points) Find $f_W(0)$.

- (e) (2 points) Find $f_{X|W}(\frac{1}{4}|W = 0)$.

- (f) (2 points) Find $P(X = \frac{2}{4}|W = 0)$.

- (g) (2 points) Find $P(X = \frac{3}{4}|W = 0)$.

- (h) (*5 bonus points*) If we don't know what value X has taken, but we can observe the value of W , what values can W take that would make us believe it is more likely that $X = \frac{2}{4}$ than either of the other two values of X ?

STAT 105 Exam II

Reference Sheet

Factorial Analysis (Two Factors)

Assuming

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

$$y_{..} = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K y_{ijk} \qquad \bar{y}_{..} = \frac{1}{n} \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K y_{ijk}$$

$$\bar{y}_{i.} = \frac{1}{n_{i.}} \sum_{j=1}^J \sum_{k=1}^K y_{ijk} \qquad \bar{y}_{.j} = \frac{1}{n_{.j}} \sum_{i=1}^I \sum_{k=1}^K y_{ijk}$$

Main effect of Factor A at level i $a_i = \bar{y}_{i.} - \bar{y}_{..}$

Main effect of Factor B at level j $b_j = \bar{y}_{.j} - \bar{y}_{..}$

Fitted Value $\hat{y}_{ij} = a_i + b_j + \bar{y}_{..}$

Discrete Random Variables

Probability function $P[X = x] = f_X(x)$

Cumulative probability function $P[X \leq x] = F_X(x)$

Expected Value $\mu = E(X) = \sum_x x f_X(x)$

Variance $\sigma^2 = Var(X) = \sum_x (x - \mu)^2 f_X(x)$

Standard Deviation $\sigma = \sqrt{Var(X)}$

Geometric Random Variables

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

Possible Values $x = 1, 2, 3, \dots$

Probability function $P[X = x] = f_X(x) = p^x(1 - p)^{x-1}$

Expected Value $\mu = E(X) = \frac{1}{p}$

Variance $\sigma^2 = Var(X) = \frac{1-p}{p^2}$

Joint Distributions and Related Distributions

Joint Probability Function $P[X = x, Y = y] = f(x, y)$

Marginal Probability Function $P[X = x] = f_X(x) = \sum_{\text{all } y} f(x, y)$
 $P[Y = y] = f_Y(y) = \sum_{\text{all } x} f(x, y)$

Conditional Probability Function $P[X = x|Y = y] = \frac{f(x, y)}{f_Y(y)}$
 $P[Y = y|X = x] = \frac{f(x, y)}{f_X(x)}$

Binomial Random Variables

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

Possible Values $x = 0, 1, 2, \dots, n$

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 $\sigma^2 = Var(X) = np(1 - p)$

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 $\sigma = \sqrt{Var(X)}$

Normal Random Variables

Let X be a normal random variable with mean μ and variance σ^2 .

Probability density function $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

Expected Value $E(X) = \mu$

Variance $Var(X) = \sigma^2$

Standard Normal Random Variables (Z)

A normal random variable with mean 0 and variance σ^2 .
If X is normal(μ, σ^2) then $P[a \leq X \leq b] = P\left[\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right]$

Probability density function $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$

Standard Normal Probabilities

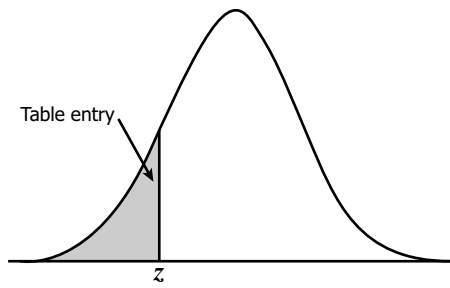


Table entry for z is the area under the standard normal curve to the left of z .

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Standard Normal Probabilities

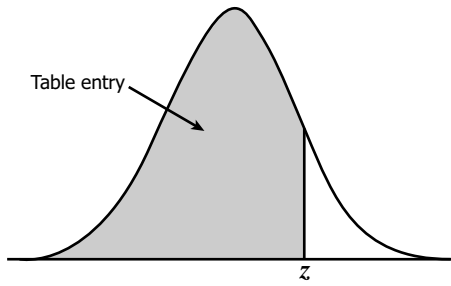


Table entry for z is the area under the standard normal curve to the left of z .

[illegible]