Stat 140, Spring 2019 **Prof. Adhikari**

Final Exam

NAME: ____ **SID:** ______________

SEAT NUMBER: ______________ **ROOM NUMBER:** ______________

You have three hours. The exam is worth 100 points. Questions 1-8 are worth 5 points each, and Questions 9-14 are worth 10 points each.

A two-sided reference sheet is provided with the exam.

INSTRUCTIONS

1. We will be scanning your exam and using Gradescope. Please write only within the black borders of each page, and write your name in the space provided on each page. Extra space for the problems on each page is provided on the reverse of the page.

2. Provide calculations or reasoning in every part of every question.

3. Unless the question says otherwise, you may leave answers as unsimplified numerical formulas including finite sums, exponential and trigonometric functions, and in terms of the standard normal c.d.f. Φ and the gamma function Γ. But **do not leave integrals or infinite sums or "..." in your answers**.

4. Some problems ask you to identify a distribution as one of the famous ones. Those distributions are listed on the reference sheet. When you name such a distribution you must also provide the parameters in the context of the problem. Note that if a problem asks you to provide "the parameters," plural, that includes the case where the distribution has just one parameter.

5. You are expected to do your own unaided work. Please read and sign the statement below.

I agree to abide by the Berkeley Honor Code. On my honor, I have neither given nor received any assistance in the taking of this exam.

Your signature: ____

This page intentionally left blank. It will not be graded.

Name: ____

1. [**5 points**] A fair die is rolled 10 times. Given that the face with six spots appears exactly two times, what is the chance that it appears on Roll 1 and Roll 10?

2. [**5 points**] Customers arrive at a bank according to a Poisson process with a rate of 20 customers per hour. Each customer withdraws cash with chance 0.8, independently of other customers. Find the chance that no more than 30 customers withdraw cash between 2 p.m. and 4 p.m.

Extra space for PROBLEMS 1 AND 2 ONLY

You may use this space for scratch work, but **only for the problems specified above**. If you need it to complete your answers on the other side of this sheet, please indicate clearly which problem you are completing, by labeling each answer with the corresponding problem number.

Name: ____

3. [**5 points**] Let $X_1, X_2, \ldots$ be i.i.d. random variables. Let $E(X_1) = \mu$ and $Var(X_1) = \sigma^2$, and for each $n \geq 1$ let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. For each $n \geq 1$, find a constant c_n such that $P(|\bar{X}_n - \mu| < c_n)$ is at least 0.99.

4. [**5 points**] Let X and Y have joint density given by

$$f(x,y) \;=\; \begin{cases} 3e^{-(2x+y)} & \text{for } 0 < x < y \\ 0 & \text{otherwise} \end{cases}$$

Fill in the first blank below with either "independent" or "not independent", and the second blank with the name and parameters of one of the famous distributions.

X and Y are ______________________, and the marginal density of X is ________________________________.

Extra space for PROBLEMS 3 AND 4 ONLY

You may use this space for scratch work, but **only for the problems specified above**. If you need it to complete your answers on the other side of this sheet, please indicate clearly which problem you are completing, by labeling each answer with the corresponding problem number.

Name: ____

5. [**5 points**] The code below results in an empirical histogram of a random variable X. Sketch the shape of the histogram; just draw a rough curve and don't worry about rectangular bars. Compute the numerical values of $E(X)$, $E(X) - SD(X)$, and $E(X) + SD(X)$, and display them on the horizontal axis as accurately as you can by eye.

```
values = make_array()
for i in np.arange(10000):
    values = np.append( values,  np.sort(stats.expon.rvs(scale = 1, size = 10)).item(0) )
Table().with_column('X', values).hist(0, bins = 25)
```

Note that `np.sort` sorts in ascending order. The `bins = 25` option to `hist` results in 25 bars, but don't worry about that for your sketch.

6. [**5 points**] Let $X_1, X_2, \ldots$ be i.i.d. exponential (1) random variables. For $n \geq 1$ let $S_n = X_1 + X_2 + \cdots + X_n$. For each probability below, fill in the first blank with either "exactly" or "approximately", and the second blank with a numerical expression.

$P(S_1 > 1)$ is ____________ equal to _________, and $P(S_{1000} > 1000)$ is ____________ equal to _________.

Extra space for PROBLEMS 5 AND 6 ONLY

You may use this space for scratch work, but **only for the problems specified above**. If you need it to complete your answers on the other side of this sheet, please indicate clearly which problem you are completing, by labeling each answer with the corresponding problem number.

Name: ____

7. [**5 points**] Consider the multiple regression model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ where $\mathbf{Y}$ is an $n \times 1$ response vector, $\mathbf{X}$ is an $n \times p$ matrix of observed non-random attributes, $\boldsymbol{\beta}$ is a $p \times 1$ vector of non-random coefficients, and $\boldsymbol{\epsilon}$ is an $n \times 1$ multivariate normal vector with mean vector $\mathbf{0}$ and covariance matrix $\sigma^2 \mathbf{I}_n$ where σ is a constant and $\mathbf{I}_n$ is the $n \times n$ identity matrix.

Let $\hat{\boldsymbol{\beta}}$ be the least squares estimate of $\boldsymbol{\beta}$, and let $\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ be the vector of predicted values of the response. Find the distribution of $\hat{\mathbf{Y}}$.

8. [**5 points**] Let X have the gamma (r, λ) distribution and assume $r > 2$. Let $V = \dfrac{X^3 + 1}{X^2}$. Find $E(V)$. **Simplify your answer so that it does not involve the gamma function.**

Extra space for PROBLEMS 7 AND 8 ONLY

You may use this space for scratch work, but **only for the problems specified above**. If you need it to complete your answers on the other side of this sheet, please indicate clearly which problem you are completing, by labeling each answer with the corresponding problem number.

Name: ____

9. A standard deck consists of four cards in each of 13 ranks, for a total of 52 cards.

A poker hand of 5 cards is dealt at random without replacement from a standard deck. Let X be the number of ranks that appear exactly once in the hand.

a) [**3 points**] Find $E(X)$.

b) [**5 points**] Find $Var(X)$.

c) [**2 points**] Find $P(X = 3)$.

Extra space for PROBLEM 9 ONLY

You may use this space for scratch work, but **only for the problem specified above**. If you need it to complete your answers on the other side of this sheet, please indicate clearly which parts you are completing, by labeling each answer with the letter corresponding to the appropriate part of the question.

Name: ____

10. Two test scores S_1 and S_2 have a bivariate normal distribution. Let $T = S_1 + S_2$. The following expectations and standard deviations are known.

- $E(S_1) = 65 = E(S_2)$
- $SD(S_1) = 10 = SD(S_2)$
- $SD(T) = 16$

a) [2 points] Let ρ be the correlation between S_1 and S_2. Find ρ. Remember to leave it as a numerical expression. You can just refer to it as ρ if you need it in the subsequent parts.

b) [2 points] Find the least squares predictor of T based on S_1.

c) [3 points] Let $D = S_1 - S_2$ and $\bar{S} = T/2$. Find the joint distribution of D and $\bar{S}$.

d) [3 points] Find $P(|D| < 7, \bar{S} > 70)$.

Extra space for PROBLEM 10 ONLY

You may use this space for scratch work, but **only for the problem specified above**. If you need it to complete your answers on the other side of this sheet, please indicate clearly which parts you are completing, by labeling each answer with the letter corresponding to the appropriate part of the question.

Name: ____

11. Let X have the beta (r, s) distribution. Given $X = p$, let $I_1, I_2, \ldots$ be i.i.d. Bernoulli (p) random variables, and for $n \geq 1$ let $S_n = I_1 + I_2 + \cdots + I_n$. Thus given $X = p$ the random variable S_n is the number of heads in n tosses of a coin that lands heads with chance p.

a) [**2 points**] For each $n \geq 1$, find $E(S_n)$.

b) [**5 points**] For each $n \geq 1$, find $Var(S_n)$.

c) [**3 points**] Let $n \geq 1$ and let $0 \leq k \leq n$. Let $m \geq 1$. Find $E(S_{n+m} \mid S_n = k)$.

Extra space for PROBLEM 11 ONLY

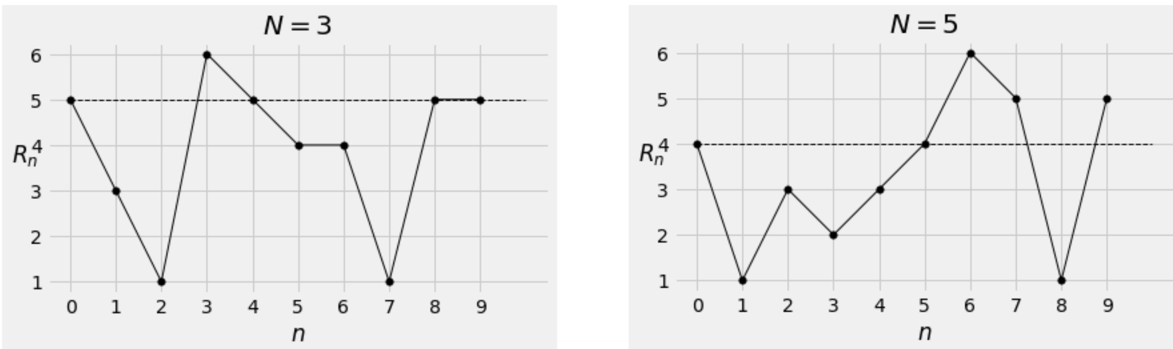
You may use this space for scratch work, but **only for the problem specified above**. If you need it to complete your answers on the other side of this sheet, please indicate clearly which parts you are completing, by labeling each answer with the letter corresponding to the appropriate part of the question.

Name: ______________________________

12. Let $R_0, R_1, R_2, \ldots$ be the numbers that appear on repeated rolls of a fair die; here "number" means "number of spots". Let N be the first time you see a number that is equal to or greater than R_0. That is,

$$N = \min\{n \geq 1 : R_n \geq R_0\}$$

For concreteness, the two figures below show examples of observed sequences $R_0, R_1, \ldots$ and the corresponding values of N.



a) [5 points] Find $E(N)$.

b) [5 points] Find the distribution of N.

Extra space for PROBLEM 12 ONLY

You may use this space for scratch work, but **only for the problem specified above**. If you need it to complete your answers on the other side of this sheet, please indicate clearly which parts you are completing, by labeling each answer with the letter corresponding to the appropriate part of the question.

Name: ____________________

13. Let X have the centered bilateral exponential density with parameter $\lambda > 0$, defined by

$$f(x) \;=\; \frac{\lambda}{2}e^{-\lambda|x|}, \qquad -\infty < x < \infty$$

a) [4 points] Find the distribution of $|X|$. It might help to sketch the graph of f.

b) [6 points] Show that the moment generating function of X is $M_X(t) = \dfrac{\lambda^2}{\lambda^2 - t^2}$ where $|t| < \lambda$.

Extra space for PROBLEM 13 ONLY

You may use this space for scratch work, but **only for the problem specified above**. If you need it to complete your answers on the other side of this sheet, please indicate clearly which parts you are completing, by labeling each answer with the letter corresponding to the appropriate part of the question.

Name: ______________________________

14. For $1 \leq i \leq 10$ let Z_i be i.i.d. standard normal random variables, and let $R = \sqrt{\sum_{i=1}^{10} Z_i^2}$.

a) [2 points] What is the distribution of R^2?

b) [4 points] Find the density of R.

c) [4 points] For $r > 0$, find $P(R > r)$.

Extra space for PROBLEM 14 ONLY

You may use this space for scratch work, but **only for the problem specified above**. If you need it to complete your answers on the other side of this sheet, please indicate clearly which parts you are completing, by labeling each answer with the letter corresponding to the appropriate part of the question.