

**V1**

Name: \_\_\_\_\_

PUID \_\_\_\_\_

Instructor (circle one): Heekyung Ahn   Evidence Matangi   Timothy Reese   Halin Shin

Class Start Time: ☐ 10:30 AM   ☐ 11:30 AM   ☐ 12:30 PM   ☐ 1:30 PM   ☐ 2:30 PM   ☐ 3:30 PM   ☐ Online

As a boilermaker pursuing academic excellence, I pledge to be honest and true in all that I do.

Accountable together - we are Purdue.

**Instructions:**

1. **IMPORTANT** Please write your **name** and **PUID** clearly on every **odd page**.
2. **Write your work in the box. Do not run over into the next question space.**
3. The only materials that you are allowed during the exam are your **scientific calculator**, **writing utensils**, **erasers**, **your crib sheet**, and **your picture ID**. Colored scratch paper will be provided if you need more room for your answers. Please write your name at the top of that paper also.
4. The crib sheet can be a handwritten or type double-sided 8.5in x 11in sheet.
5. Keep your bag closed and cellphone stored away securely at all times during the exam.
6. If you share your calculator without permission or have a cell phone at your desk, you will get a **zero** on the exam. Do not take out your cell phone until you are next in line to submit your exam.
7. The exam is only 60 minutes long so there will be no breaks during the exam. If you leave the exam room, you must turn in your exam, and you will not be allowed to come back.
8. **For free response questions you must show ALL your work to obtain full credit.** An answer without showing any work may result in **zero** credit. If your work is not readable, it will be marked wrong. Remember that work has to be shown for all numbers that are not provided in the problem or no credit will be given for them. All explanations must be in complete English sentences to receive full credit.
9. All numeric answers should have **four decimal places** unless stated otherwise.
10. After you complete the exam, please turn in your exam as well as your table and any scrap paper that you used. Please be prepared to **show your Purdue picture ID**. You will need to **sign a sheet** indicating that you have turned in your exam.
11. You are expected to uphold the honor code of Purdue University. It is your responsibility to keep your work covered at all times. Anyone caught cheating on the exam will automatically fail the course and will be reported to the Office of the Dean of Students.
12. It is strictly prohibited to smuggle this exam outside. Your exam will be returned to you on Gradescope after it is graded.

**Your exam is not valid without your signature below. This means that it won't be graded.**

I attest here that I have read and followed the instructions above honestly while taking this exam and that the work submitted is my own, produced without assistance from books, other people (including other students in this class), notes other than my own crib sheet(s), or other aids. In addition, I agree that if I tell any other student in this class anything about the exam BEFORE they take it, I (and the student that I communicate the information to) will fail the course and be reported to the Office of the Dean of Students for Academic Dishonesty.

Signature of Student: \_\_\_\_\_

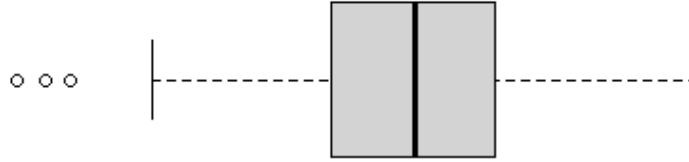
**You may use this page as scratch paper.  
The following is for your benefit only.**

<b>Question Number</b>	<b>Total Possible</b>	<b>Your points</b>
Problem 1 (True/False) (2 points each)	12	
Problem 2 (Multiple Choice) (3 points each)	15	
Problem 3	26	
Problem 4	26	
Problem 5	26	
Total	105	

**The rest of this page can be used for scratch work**

1. (12 points, 2 points each) True/False Questions. Indicate the correct answer by completely filling in the appropriate circle. If you indicate your answer by any other way, you may be marked incorrect.

1.1. The boxplot below visually displays the summary information for a dataset.



☒ or ☐ According to the boxplot above, a lower inner fence must be located between (inclusive) the lower whisker and the largest outlier.

1.2. The Binomial distribution is symmetric when

☐ or ☒ the probability of success  $p$  is close to 0 or 1.

1.3. Suppose  $Y$  is the outcome of a single roll of a 6-sided die with an unknown probability mass function having nonzero variance. The outcome for a random variable  $X$  is obtained by throwing this die once then multiplying the resulting number by 3, i.e.,  $X = 3Y$ . The outcome for another random variable  $Z$  is obtained by throwing the same die three times, then adding the results together, i.e.,  $Z = Y_1 + Y_2 + Y_3$  where  $Y_1, Y_2, Y_3$  are independent copies of  $Y$ . Then,

☐ or ☒ it follows that  $\text{Var}(X) = \text{Var}(Z)$ .

1.4. Let  $X$  denote a normal random variable, then regardless of the value of  $E[X]$  and  $\text{Var}(X)$ ,

☒ or ☐  $P(E[X] - 2\sqrt{\text{Var}(X)} < X < E[X] + 2\sqrt{\text{Var}(X)}) \approx 0.95$  is always true.

1.5. Let  $X$  be an exponential random variable. Then,

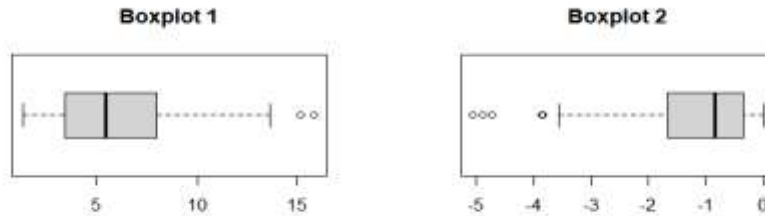
☐ or ☒ the distribution of  $X$  models the probability associated with the total number of events occurring during a fixed interval of time.

1.6. Let  $X$  be a continuous random variable with finite expected value  $\mu$  and variance  $\sigma^2$ . Define a new random variable  $Y = aX + b$  where  $a, b$  are real numbers with  $a \neq 0$ .

☒ or ☐ It follows that  $E[Y^2] = a^2(\sigma^2 + \mu^2) + 2ab\mu + b^2$  always holds, regardless of the distribution of  $X$ .

2. (15 points, 3 points each) **Multiple Choice Questions.** Indicate the correct answer by completely filling in the appropriate circle. If you indicate your answer by any other way, you may be marked incorrect. **For each question, there is only one correct option letter choice.**

2.1. Which of the following provides the best measures of center and spread respectively, based on the boxplots below?



- ☐ (A) **Boxplot 1:** sample mean, interquartile range (IQR)
- ☐ (B) **Boxplot 1:** sample median, range
- ☐ (C) **Boxplot 1:** sample mean, sample standard deviation
- ☒ (D) **Boxplot 2:** sample median, interquartile range (IQR)
- ☐ (E) **Boxplot 2:** sample median, range
- ☐ (F) **Boxplot 2:** sample mean, sample standard deviation

2.2. Assume  $X \sim N(\mu, \sigma)$ ,  $Y \sim \text{Exp}(\lambda)$ , and  $Z \sim \text{Bin}(n, p)$ . Which of the following statements is TRUE?

- ☐ (A) For the parameters of  $X$  the ratio  $\mu/\sigma$  must be greater than 1.
- ☒ (B) For any  $y$  in the support of  $Y$  there exists an  $x$  in the support of  $X$  such that  $x = y$ .
- ☐ (C) The variance of  $Y$  is the same as the mean of  $Y$ .
- ☐ (D) If  $p < 0.5$ , then  $Z$  can never take any values greater than  $n/2$ . In other words,  $Z$  is supported only on integers strictly less than  $n/2$ .
- ☐ (E) The parameter  $\lambda$  must be a positive integer.

2.3. Harry works at Hogwarts mail center. The number of owls that he receives in an hour at the center ( $X$ ) follows Poisson distribution with an average hourly rate of 1. In other words,  $X \sim \text{Poisson}(\lambda = 1)$ . Which of the followings is not true?

- ☐ (A) The probability that Harry receives  $x$  owls in an hour is  $\frac{1^x e^{-1}}{x!}$ .
- ☐ (B) The probability that Harry receives 1 owl in an hour is approximately 0.3679.
- ☐ (C) The probability that Harry receives 2 owls in an hour is approximately 0.1839.
- ☒ (D) The probability that Harry receives more than 2 owls in an hour is approximately 0.2642.
- ☐ (E) The probability that Harry receives 1 owl in the first hour of his shift then zero in the second hour is approximately 0.1353.

2.4. If  $X$  is a Poisson random variable, that satisfies  $P(X = 5) = P(X = 7)$ , then  $P(X = 0) = ?$

Ⓐ  $P(X = 0) = e^0$

Ⓑ  $P(X = 0) = e^{-5}$

Ⓒ  $P(X = 0) = e^{-7}$

Ⓓ  $P(X = 0) = e^{-\sqrt{35}}$

Ⓔ  $P(X = 0) = e^{-\sqrt{42}}$

2.5. In the **standard Normal distribution**, for any  $z > 0$ , how does the probability compare between the two regions  $-z < Z < 0$  and  $0 < Z < z$ ? Determine the correct symbol connecting the two probability statements below (fill in the blank).

$$P(-z < Z < 0) \text{ \_\_\_\_\_\_ } P(0 < Z < z)$$

Ⓐ  $<$

Ⓑ  $>$

Ⓒ  $=$

Ⓓ  $\neq$

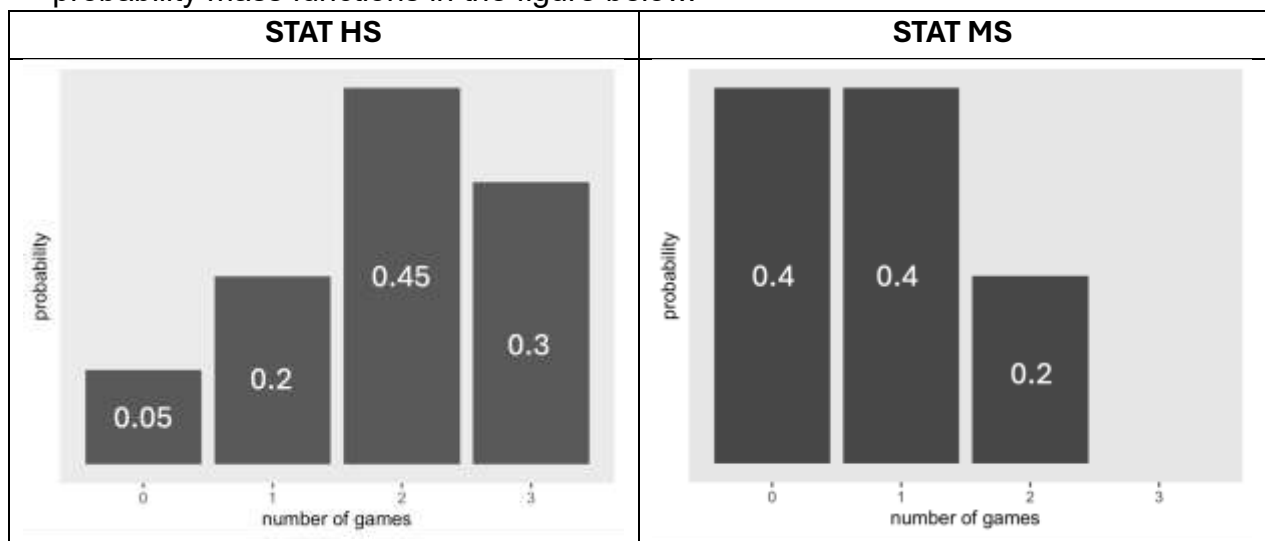
Ⓔ  $\subset$

**Free Response Questions 3-5.** Show all work, clearly label your answers, and use **four decimal places**.

3. **(26 points)** In an alternative timeline where Lafayette evolved into a structured statistical metropolis, two competing academic institutions, STAT High School (**STAT HS**) and STAT Middle School (**STAT MS**), compete for control of a shared soccer field used for team practices and school games.

The field is reserved on a monthly basis, with STAT High School holding the reservation **74% of the time**, independent of other months. Whenever STAT High School does not secure the reservation, it is automatically assigned to STAT Middle School.

The number of games played each month follows the distribution depicted by the probability mass functions in the figure below.



- a) **(6 points)** Given that **STAT HS** will hold **at least one** soccer game this month, what is the probability that it holds **exactly three**?

Let  $N_G$  denote the number of games held during a given month it can take values in the set  $\{0, 1, 2, 3\}$ .

$$\begin{aligned}
 &P(N_G = 3 | \{N_G \geq 1\} \cap \{\text{STAT HS}\}) \\
 &= \frac{P(\{N_G = 3\} \cap \{\text{STAT HS}\})}{P(\{N_G \geq 1\} \cap \{\text{STAT HS}\})} \\
 &= \frac{P(N_G = 3 | \text{STAT HS})P(\text{STAT HS})}{P(N_G \geq 1 | \text{STAT HS})P(\text{STAT HS})} \\
 &= \frac{0.3}{0.95} = 0.3158
 \end{aligned}$$

- b) (6 points) On any given month, what is the probability that **at least one game is played**?

$$\begin{aligned}
 P(N_G \geq 1) &= P(N_G \geq 1 | \text{STAT HS})P(\text{STAT HS}) + P(N_G \geq 1 | \text{STAT MS})P(\text{STAT MS}) \\
 &= 0.95 \times 0.74 + 0.6 \times (1 - 0.74) \\
 &= 0.859
 \end{aligned}$$

- c) (8 points) Knowing that **at least one game is held next month**, what is the probability that the reservation is held by **STAT MS**?

$$\begin{aligned}
 P(\text{STAT MS} | N_G \geq 1) &= \frac{P(\{\text{STAT MS}\} \cap \{N_G \geq 1\})}{P(N_G \geq 1)} \\
 &= \frac{P(N_G \geq 1 | \text{STAT MS})P(\text{STAT MS})}{0.859} \\
 &= \frac{0.6 \times (1 - 0.74)}{0.859} \\
 &= \frac{0.156}{0.859} \\
 &= 0.1816
 \end{aligned}$$

- d) (6 points) Are the **reservation holder** and the **number of games played** in a month **independent**? Justify your conclusion mathematically.

No the number of games played in a given month is not independent of who holds the reservation. We can either use the results from part c) to see this:

$$P(\text{STAT MS} | N_G \geq 1) = 0.156 \neq 0.26 = P(\text{STAT MS})$$

Or a more obvious way to draw this conclusion is that if STAT MS holds the reservation, we know for a fact that 3 games will not be played and hence the number of games played is not independent of who holds the reservation.

4. (26 points) A zombie enthusiast is studying the walking speeds of classic zombies. Based on extensive observations, the enthusiast concludes that the speed of a classic zombie follows a normal distribution with:

- a) Mean  $\mu = 2$  miles per hour (mph)
- b) Standard Deviation  $\sigma = 0.19$  mph

Use this information to address the following questions:

- a) (4 points) What is the probability that a randomly chosen classic zombie walks **faster than 2.3 mph**?

Let  $F$  denote the distribution of how fast a classic zombie walks then we have that

$$F \sim N(\mu = 2, \sigma = 0.19).$$

$$P(F > 2.3) = P\left(Z > \frac{2.3 - 2}{0.19}\right) = P(Z > 1.58) = P(Z < -1.58) = 0.0571$$

- b) (6 points) Given that a randomly chosen classic zombie walks **faster than 2 mph**, what is the probability that it also walks **faster than 2.3 mph**?

$$P(F > 2.3 | F > 2) = \frac{P(F > 2.3)}{P(F > 2)} = \frac{0.0571}{0.5} = 0.1142$$

- c) (6 points) If the enthusiast **randomly selects 10** classic zombies, what is the probability that **at least one** has a speed **greater than 2.3 mph**

Let  $F_1, F_2, \dots, F_{10}$  represent the random variables for 10 class zombies we can assume that they all are independent and follow the same distribution

$$F_i \sim N(\mu = 2, \sigma = 0.19) \text{ for } i = 1, 2, \dots, 10.$$

$$\begin{aligned} &P(\{F_1 \geq 2.3\} \cup \{F_2 \geq 2.3\} \cup \dots \cup \{F_{10} \geq 2.3\}) \\ &= P(\{F_1 < 2.3\} \cap \{F_2 < 2.3\} \cap \dots \cap \{F_{10} < 2.3\})' \\ &= 1 - P(F_1 < 2.3)^{10} \\ &= 1 - (1 - P(F > 2.3))^{10} \\ &= 1 - (1 - 0.0571)^{10} \\ &= 1 - 0.9429^{10} \\ &0.4445 \end{aligned}$$



- d) (4 points) Suppose **two** classic zombies both begin traveling at the same time. After **3 hours**, what is the **expected total distance** they will have covered **combined**?

Since both classic zombies follow the same distribution with an expected walking speed of 2 miles per hour and they both walk for 3 hours we can conclude after 3 hours each zombie will have covered an average of 6 miles for a total of 12 miles combined.

- e) (6 points) A classic zombie is considered **Elite** if its speed is in the **top 3%**. Determine the **minimum** speed at which a classic zombie is classified as Elite.

Top 3% walking speed is the transformed z score for the bottom 97<sup>th</sup> percentile.

$$z_{0.97} = 1.88$$

Let  $f^*$  denote the top 2% walking speed then we have

$$f^* = 2 + 0.19 \cdot 1.88 = 2.3572$$

5. (26 points) At a busy taco truck, customers wait different amounts of time depending on order complexity and queue length. Let  $X$  be the total time, in minutes, from joining the line to receiving food.

No one is served in less than five minutes. The likelihood of finishing follows a parabolic pattern between five and seven and a half minutes, increasing until it peaks at seven and a half. After that, the rate of completion remains constant until twelve and a half minutes, when all orders are fulfilled.

The proposed probability density function has the form

$$f_X(x) = \begin{cases} k \cdot (6.25 - (x - 7.5)^2) & 5 \leq x < 7.5 \\ \frac{25}{4} \cdot k & 7.5 \leq x < 12.5 \\ 0 & \text{otherwise} \end{cases}$$

where  $k$  must be determined so that  $f_X(x)$  is a valid probability density function.

- a) (18 points) Determine the value of  $k$  such that the function  $f_X(x)$  is a valid PDF.

$$k \left[ \int_5^{7.5} 6.25 dx - \int_5^{7.5} (x - 7.5)^2 dx \right] + \frac{25}{4} k \int_{7.5}^{12.5} dx = 1$$

$$k \left[ \frac{125}{8} - \int_{-2.5}^0 u^2 du \right] + \frac{125}{4} k = 1$$

$$k \left[ \frac{125}{8} - \frac{125}{24} \right] + \frac{125}{4} k = 1$$

$$= \frac{250}{24} k + \frac{750}{24} k = 1$$

$$k = \frac{24}{1000} = \frac{3}{125} = 0.024$$

The cumulative distribution function for the total time, in minutes, from joining the line to receiving food is given by:

$$F_X(x) = \begin{cases} 0 & x < 5 \\ -\frac{1}{125}(x - 7.5)^3 + \frac{3}{20}x - \frac{7}{8} & 5 \leq x < 7.5 \\ \frac{3}{20}x - \frac{7}{8} & 7.5 \leq x < 12.5 \\ 1 & x \geq 12.5 \end{cases}$$

- b) (8 points)** The taco truck owner wants to know how long a typical customer waits before receiving their order. Instead of looking at the average, they are interested in the median wait time, the time by which half of all customers have received their food.

Determine the median wait time  $\tilde{\mu}$ , where  
 $P(X \leq \tilde{\mu}) = 0.5$ .

From the work in part a) we know that the first region accumulates a total area of

$$k \left[ \int_5^{7.5} 6.25 dx - \int_5^{7.5} (x - 7.5)^2 dx \right] = \frac{250}{24} k = 0.25$$

Therefore the 50<sup>th</sup> percentile (median) must occur in region 3:

$$\frac{3}{20}x_{0.5} - \frac{7}{8} = \frac{1}{2}$$

$$\frac{3}{20}x_{0.5} = \frac{11}{8}$$

$$x_{0.5} = \frac{11}{8} \times \frac{20}{3}$$

$$x_{0.5} = 9.1667$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

