Your Name: ___________________________________________

Your Instructor: Jeff Tianhong Jeremy Yong
Yen-Ning Mike Zhaonan

Your class time (circle one):

<table>
<thead>
<tr>
<th>7:30</th>
<th>8:30</th>
<th>9:30</th>
<th>10:30</th>
<th>11:30</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12:30</td>
<td>1:30</td>
<td>2:30</td>
<td>3:30</td>
<td>4:30</td>
</tr>
</tbody>
</table>

Note:

- Show your work on all questions. Unsupported work will not receive full credit.
- All answers should be in decimal form and should be exact, or to at least taken out to two decimal places.
- You are responsible for upholding the Honor Code of Purdue University. This includes protecting your work from other students.
- You are allowed the following aids: a one-page 8 ½” x 11” handwritten cheat sheet (in your handwriting), a scientific calculator, and pencils.
- Instructors will not interpret questions for you. If you do have questions, wait until you have looked over the whole exam so that you can ask all of your questions at one time.
- You must show your student ID (upon request) and turn in your cheat sheet when you turn in your exam to your instructor.
- Turn off your cell phone before the exam begins!

<table>
<thead>
<tr>
<th>Question</th>
<th>Points Possible</th>
<th>Points Received</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Cheat Sheet</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>
1. In each of the following situations, there is a random variable involved. Name its exact distribution and find the parameter(s). If an approximation can be made, you must name and label both the exact and approximate distribution to receive full credit. In cases where only an approximation is possible to your knowledge, you only need to state the name and parameter(s) of the approximating distribution. (3 points each)

a). Someone is using Windows 98™ on his laptop. When the system freezes, it takes anywhere between 3 to 5 minutes for it to get back to work. Let $F$ represent the amount of system freezing time.

$$F\sim\text{Uni}(3, 5) \text{ Exact}$$

b). A quarterback on a football team has a passing completion percentage of 64% throughout his career. He attempted 45 passes during a given game. Let $C$ represent the number of completed passes he made in this game.

Exact: $C\sim\text{BIN}(45, 0.64)$

Approx: $C\sim\text{Normal}(45\times0.64=28.8, 45\times0.64\times0.36=10.368)$

c). There are an average of 3.5 gas stations on every 100 miles of interstate highway. Let $D$ be the distance between gas stations on interstate highways.

$$D\sim\text{Exp}(100/3.5) \text{ Exact}$$

d). 25% of all the new $100 bills have printing problems. You just cashed a one million dollar check from bank, all in $100 bills. Let $P$ be the proportion of these bills that have printing problems.

$$P\sim\text{Normal}(0.25, 0.0001875) \text{ Approximate}$$

e). There are 10 quizzes in a stat course. For each quiz, the course average follows a normal distribution with mean of 16.75 and variance of 25. Let $A$ represent the average of the 10 course averages.

$$A\sim\text{Normal}(16.75, 2.5)$$
2. Let $T$ denote the amount of time a personal banker spends with each customer, and suppose $T$ follows an exponential distribution with an average amount of time equal to 25 minutes.

a). What is the probability that a personal banker spends less than 20 minutes with one customer? (2 points)

$$P(T<20)=1-\exp(-20/25)=0.55$$

b). Given a personal banker has spent 15 minutes on one customer, what is the probability that the banker will spend less than 25 minutes with this customer? (3 points)

$$P(T<25 \mid T >15)= \frac{P(15<T<25)}{P(T>15)}=0.40$$

c). What is the median of $T$? (3 points)

$$T_m=-25\ln(0.5)=17.33$$

d). It is believed that the energy, $N$, a personal banker spends in the process of dealing with a customer is $N=mT^2$, where $m$ is the weight of the banker in pounds and $T$ is the amount of time spent. For a banker who weighs 180 lb, what will the expected amount of energy he spends with each customer? (3 points)

Let $N$ be the amount of energy, then $E(N)=180E(T^2)$

$E(T)=25$, $E(T^3)=\text{var}(T)+E^2(T)=25^2+25^2=1250$

Therefore, $E(N)+180*1250=225000$

e). What is the approximate probability that the personal banker can spend less than 20 minutes with at least 15 out of the next 30 customers. (Hint: you can use the result from part a) (3 points)

Let $K$ be the number of customers that the banker spends less than 20 minutes with, then $K \sim \text{BIN}(30, 0.55)$

We can use normal to approximate $K$ by $K \sim \text{Normal}(30*0.55, 30*0.55*0.45)$

The approximate probability is $P(K\geq15)=P(K>14.5)=1-P(K<14.5)==1-P(Z<-0.73)=0.2315$
3. Let X be a continuous random variable with PDF:

\[ f(x) = \begin{cases} 
C(3x^2 - 18x + 27), & 1 < x < 3 \\
0, & \text{otherwise}
\end{cases} \]

where C is a constant.

a). Find the value of C such that f(x) is a valid PDF. (3 points)

\[
\int_1^3 f(x)dx = 1 \Rightarrow \int_1^3 C(3x^2 - 18x + 27)dx = 1 \Rightarrow C[x^3 - 9x^2 + 27x]_1^3 = 1 \Rightarrow 8C = 1 \Rightarrow C = 1/8
\]

b). Find the CDF of X. (3 points)

\[
F(X) = \int_1^x f(x)dx \Rightarrow (1/8)\int_1^x (3x^2 - 18x + 27)dx \Rightarrow (1/8)[x^3 - 9x^2 + 27x]_1^x \Rightarrow (1/8)(x^3 - 9x^2 + 27x - 19)
\]

Therefore, \(F(X) = \begin{cases} 
0, & \text{if } X \leq 1 \\
(1/8)(x^3 - 9x^2 + 27x - 19), & \text{if } 1 < X < 3 \\
1, & \text{if } X \geq 3
\end{cases}\)

c). Find the 87.5\(^{th}\) percentile of X. (3 points)

\[
F(X) = 0.875, \text{ then } (1/8)(x^3 - 9x^2 + 27x - 19) = 0.875 \\
\Rightarrow x^3 - 9x^2 + 27x - 26 = 0 \\
\Rightarrow (x - 2)(x^2 - 7x + 13) = 0 \\
\Rightarrow X = 2
\]
Problem 3. (contd)

d). Find $P( X < 2.5|X > 1.5)$. (3 points)

$$P( X < 2.5|X > 1.5)=P(1.5<X<2.5) / P(X>1.5) = [F(2.5)-F(1.5)]/[1-F(1.5)]=0.963$$

e). Find $E[3X]$ (4 points)

$$E(3X)=\int_{1}^{3} 3x*f(x)dx$$
$$\rightarrow (3/8)\int_{1}^{3} (3x^3-18x^2+27x)dx$$
$$\rightarrow (3/8)[(3/4)x^4-6x^3+(27/2)x^2]_{1}^{3}$$
$$\rightarrow (3/8)*12=4.5$$
4. The highway speed of all trucks has a standard deviation is 12.5 mph. A random sample of 120 trucks shows a mean highway speed of 62 mph.

a). Construct a 90% confidence interval for the mean highway speed of the population. (3 points)

\[(62 - 1.645 \times 12.5 / \sqrt{120}, 62 + 1.645 \times 12.5 / \sqrt{120})\]
\[= (60.12, 63.88)\]

b). If we want to be 95% confident that the margin of error of the sample mean is less than 2 mph, how many more trucks do we need to sample? (3 points)

\[n = \left(\frac{1.96 \times 12.5}{2}\right)^2 = 150.0625 \rightarrow 151\]

Therefore, we need to sample 151-120=31 more trucks.

c). In this sample of 120 trucks, 48 trucks were driving at a speed higher than 65 mph, which means they are speeding. Construct an 80% confidence interval for the proportion of the trucks that are speeding among all trucks driving on the highway. (3 points)

\[(\frac{48}{120} - 1.285 \times \sqrt{0.4 \times 0.6 / 120}, \frac{48}{120} + 1.285 \times \sqrt{0.4 \times 0.6 / 120})\]
\[= (0.4 - 0.057, 0.4 + 0.057)\]
\[= (0.343, 0.457)\]
5. Lost Hills Food Inc. sells their famous “Delicioso” cookies in small boxes. The weight of each individual box of cookies follows an unknown distribution with mean of 35 ounces and variance of 64.

a). You just purchased 100 boxes of “Delicioso” for a party and you are interested in the average weight, $W$, of these boxes. Name an appropriate distribution and its parameter(s) for $W$. (3 points)

$$W \sim \text{Normal}(35, 0.64)$$

b). Find the probability that the average weight of the cookies in a box exceeds 36 ounces. (3 points)

$$P(W > 36) = 1 - P(W < 36) = 1 - P(Z < 1.25) = 0.1056$$

c). The company’s production records show that 2.5% of all the boxes of “Delicioso” cookies are “under-filled”, which means their weight is below a given amount. What would be that amount? (3 points)

$$W_{2.5} = 35 - 1.96 \times 0.8 = 33.432$$
6. A biologist set up 50 identical traps to collect ants for a study. Her knowledge tells her that the number of ants collected by each trap follows a Poisson distribution with a mean of 2.2. However, she needs a trap to collect at least 3 ants to be useful. (3 points each)

a). What is the probability that a trap will collect enough ants for this biologist’s study?

\[ P(X\geq 3) = 1 - P(X=0) - P(X=1) - P(X=2) = 0.3773 \]

b). Let \( T \) represent the number of traps that will collect enough ants for the study, name the exact distribution and the parameter(s) for \( T \).

\[ T \sim \text{BIN}(50, 0.3773) \]

c). Find the exact probability that 30 traps will collect enough ants for the study.

\[ P(T=30) = \binom{50}{30} 0.3773^{30} (1-0.3773)^{20} = 0.0007 \]

d). Name a good approximating distribution for \( T \) and its parameter(s).

\[ T \sim \text{Normal}(50 \times 0.3773 = 18.865, 50 \times 0.3773 \times 0.6227 = 11.747) \]

e). The biologist needs at least 25 of her traps to collect enough ants for the study. What is the approximate probability that she can get enough useful traps?

\[ P(T\geq 25) = 1 - P(T<24.5) = 1 - P(1.64) = 0.05 \]
7. Two brothers, Larry and Harry, are both former high school track athletes. Let L and H denote the time in minutes it takes each brother to run 1 mile, respectively. Assume these times are independent and follow the distributions given below:

\[ L \sim \text{Unif}(5, 7) \text{ and } H \sim \text{Unif}(6, 8) \]

a). Compute the probability it takes Larry less than 5.5 minutes to run 1 mile. (3 points)

\[ P(L < 5.5) = (5.5 - 5)/(7 - 5) = 0.4 \]

b). Seven minutes ago Harry began running 1 mile and he is not yet done. Taking this information into consideration, find the probability that it takes him at least 30 more seconds to finish. (3 points)

\[ P(H > 7.5 \setminus H > 7) = P(H > 7.5)/P(H > 7) = (8 - 7.5)/(8 - 7) = 0.5 \]

c). What is the probability that at least one of the brothers finishes within 6.5 minutes? (4 points)

\[
\begin{align*}
P(\text{neither finishes within 6.5 seconds}) &= 1 - P(L > 6.5 \text{ and } H > 6.5) \\
&= 1 - P(L > 6.5) \cdot P(H > 6.5) \\
&= 1 - 0.5/2 \cdot 1.5/2 \\
&= 0.8125
\end{align*}
\]
Once upon a time, on a snowy day in December, Snow White and her seven dwarfs paid a visit to Santa Claus. Her dwarfs were so excited to see Santa’s eight reindeers in the house and they all wanted to ride on them.

(a). Assuming each dwarf can only ride on one reindeer and no two dwarfs should ride on the same reindeer, how many possible ways are there to arrange the rides? (2 points)

\[ P_8^7 = 40320 \]

(b). Cheered by the curiosity of the dwarfs, Santa brought them to his farm where 100 reindeers were raised. Among all the reindeers on the farm, 20 of them have a red nose. Santa selected seven different reindeers for the dwarfs to ride on. What is the probability that 3 red-nosed reindeer were selected? (3 points)

Let R be the number of red-rosed reindeers, \( R \sim HG(100, 7, 20) \)

\[ P(R=3) = \frac{C_{20}^3 \cdot C_{80}^4}{C_{100}^7} = 0.1126 \]

c). The dwarfs fell in love with the reindeer rides and they made an annual appointment with Santa for the next 40 years. Let \( R_i \) represent the number of red-nosed reindeers that Santa selected for the dwarfs to ride on each year and let \( \bar{R} \) represent the average of all \( R_i \)'s. Assuming Santa has the same herd of reindeers on his farm all the time, name an approximating distribution for \( \bar{R} \) and its parameter(s). (3 points)

\[ \bar{R} \sim Normal(7\times0.2=1.4, 7\times0.2\times0.8\times\sqrt{\frac{100-7}{100-1}/40}=0.027) \]

d). What is the probability that, on average, the dwarfs will ride on more than 2 red-nosed reindeers during each of the 40 annual reindeer ride event? (3 points)

\[ P(\bar{R}>2) = 1 - P(\bar{R} \leq 2) = 1 - P(\bar{R} < 2.5) = 1 - P(Z < 6.78) \rightarrow \text{almost 1} \]