Fall 2014 **Statistics 330** **Exam II**

(100 points)

Please write your first and last name here:

Name ________________________________

Instructions:

- Partial credit will be given only if you show your work.
- Reason out your answers. In many cases, a line or two of justification is enough.
- The questions are roughly in the order in which the material is presented in class, so they are not necessarily ordered easiest to hardest.
- If you get stuck on one, it may be a good idea to move on and come back to that question at the end.
- You may use your prepared notes (1 page, both sides) and a calculator only.

1. Suppose X is a discrete random variable with probability distribution given in below table. Find the variance of X.

x	-2	-1	0	1	2
$p(x)$	0.25	0.1	0.3	0.1	a

   (a) What is a? (5 points)

   Answer:

$$\begin{aligned}\sum_x p(x) &= 1 \\ 0.25 + 0.1 + 0.3 + 0.1 + a &= 1 \\ 0.75 + a &= 1 \\ a &= 0.25\end{aligned}$$

   (b) What are the expectation and variance of X? (10 points)

   Answer:

$$\begin{aligned}E(X) &= \sum_x xp(x) \\ &= (-2)(0.25) + (-1)(0.1) + (0)(0.3) + (1)(0.1) + (2)(0.25) = 0 \\ Var(X) &= \sum_x (x - E(X))^2 p(X) \\ &= (-2-0)^2(0.25) + (-1-0)^2(0.1) + (0-0)^2(0.3) \\ &\quad + (1-0)^2(0.1) + (2-0)^2(0.25) \\ &= 2.2\end{aligned}$$

   (c) What is $P(X \leq 0)$? (5 points)

   Answer:

$$\begin{aligned}P(X \leq 0) &= P(X = -2) + P(X = -1) + P(X = 0) \\ &= p(-2) + p(-1) + p(0) \\ &= 0.25 + 0.1 + 0.3 = 0.65\end{aligned}$$

2. You are given a binomial random distribution with 9 trails and probability of success on a single trial being 0.45.

   (a) What is the probability of getting more than 2 successes in 9 trails? (Keep at least 4 decimal places in your calculations.) (10 points)

   Answer:

$$\begin{aligned} P(\text{more than 2 successes}) &= 1 - P(\text{less than or equal to 2 successes}) \\ &= 1 - (p(0) + p(1) + p(2)) \\ &= 1 - \binom{9}{0} 0.45^0 (1-0.45)^9 - \binom{9}{1} 0.45^1 (1-0.45)^8 \\ &\quad - \binom{9}{2} 0.45^2 (1-0.45)^7 \\ &= 1 - (0.0046 + 0.0339 + 0.1110) \\ &= 0.8505 \end{aligned}$$

   (b) What is the expectation and variance of this distribution? (5 points)

   Answer: The expectation is $np = 9 \times 0.45 = 4.05$. The variance is $np(1-p) = 9 \times 0.45 \times (1-0.45) = 2.2275$.

   (c) Instead of 9 trials, we repeat the experiment until the first success. Let W be the number of trails. For instance, if the first success happened in the second trial, $W = 2$. What are the expectation and variance of W? (5 points)

   Answer: W follows geometric distribution with $p = 0.45$. $E(W) = 1/p = 1/0.45 = 2.2222$ and $Var(W) = (1-p)/p^2 = 2.7160$.

3. Consider the following joint distribution for two random variables X and Y:

X \ Y	0	1	2	3
0	0.2	0.1	0.1	0
1	0.3	0.1	0	0.2

(a) Find the marginal distributions for X and Y and the $E[X]$ and $E[Y]$. (8 points)

Answer: The marginal distributions for X and Y are is

x	0	1
$p_X(x)$	0.4	0.6

y	0	1	2	3
$p_y(y)$	0.5	0.2	0.1	0.2

The expectations are

$$E[X] = 0 \times 0.4 + 1 \times 0.6 = 0.6$$
$$E[Y] = 0 \times 0.5 + 1 \times 0.2 + 2 \times 0.1 + 3 \times 0.2 = 1$$

(b) Find $Var[X]$, $Var[Y]$, and $\text{Cov}(X, Y)$. (10 points)

Answer:

$$\begin{aligned} V[X] =& (0 - .6)^2 \times 0.4 + (1 - 0.6)^2 \times 0.6 = 0.24 \\ V[Y] =& (0 - 1)^2 \times 0.5 + (1 - 1)^2 \times 0.2 + (2 - 1)^2 \times 0.1 + (3 - 1)^2 \times 0.2 = 1.4 \end{aligned}$$

and

$$\begin{aligned} \text{Cov}(X, Y) = & \quad (0 - 0.6)(0 - 1) \times 0.2 && +(1 - 0.6)(0 - 1) \times 0.3 \\ & + (0 - 0.6)(1 - 1) \times 0.1 && +(1 - 0.6)(1 - 1) \times 0.1 \\ & + (0 - 0.6)(2 - 1) \times 0.1 && +(1 - 0.6)(2 - 1) \times 0 \\ & + (0 - 0.6)(3 - 1) \times 0 && +(1 - 0.6)(3 - 1) \times 0.2 \\ =& 0.1 \end{aligned}$$

(c) Are X and Y independent? Why or why not? (2 points)

Answer: Since $0 = P(X = 0, Y = 3) \neq P(X = 0)P(Y = 3) = 0.4 \times 0.2 = 0.08$, they are NOT independent.

4. In an industrial setting, a hard drive has about a 15% chance of failing in a given year. Let X be the failure time in years for an individual hard drive. If we assume X has an exponential distribution, then $X \sim \text{Exp}(\lambda = 0.16)$.

   (a) What is the expected failure time for an individual hard drive? (5 points)

   Answer: $E[X] = 1/\lambda = 1/0.16 \approx 6.25$ years.

   (b) What is the probability that a hard drive lasts longer than 2 years? (5 points)

   Answer: $P(X > 5) = e^{-\lambda \times 2} = e^{-0.16 \times 2} = 0.73$

   (c) If a hard drive has already lasted 2 years, what is the probability it will last another year? (5 points)

   Answer: By the memoryless property,

$$P(X > 2 + 1 | X > 2) = P(X > 1) = e^{-\lambda \times 1} = e^{-0.16} = 0.85.$$

   (d) Given the 15% chance of first year failure, derive $\lambda = 0.16$. (5 points)

   Answer:

$$0.15 = P(X < 1) = 1 - e^{-\lambda} \implies \lambda = -\log(1 - 0.15) \approx 0.16$$

   More information about true hard drive failure rates can be found here:
   `www.extremetech.com/computing/170748-how-long-do-hard-drives-actually-live-for`

5. A continuous random variable X has the probability density function (pdf)

$$f_X(x) = \begin{cases} cx & \text{if } 2 < x < 3 \\ 0 & \text{otherwise.} \end{cases}$$

(If you cannot do parts a or b, you should still be able to do parts c and d.)

(a) Show that $c = 0.4$ makes $f_X(x)$ a valid pdf. (5 points)

Answer:

$$1 = \int_2^3 cx dx = c\frac{x^2}{2}\bigg|_2^3 = c\left[\frac{3^2}{2} - \frac{2^2}{2}\right] = c \times 2.5$$

so $c = 1/2.5 = 0.4$.

(b) Show that $E[X] = 2.533$. (5 points)

Answer:

$$E[X] = \int_2^3 x \times cx dx = c\frac{x^3}{3}\bigg|_2^3 = c\left[\frac{3^3}{3} - \frac{2^3}{3}\right] = c \times 6\frac{1}{3} \approx 2.533$$

The variance is

$$\begin{aligned} V[X] &= \int_2^3 (x-\mu)^2 \times cx dx = c\int_2^3 x^3 - 2\mu x^2 + \mu^2 x dx \\ &= c\left[\frac{x^4}{4} - 2\frac{x^3}{3} + \mu^2\frac{x^2}{2}\right]\bigg|_2^3 = c\left[\left(\frac{3^4}{4} - 2\frac{3^3}{3} + \mu^2\frac{3^2}{2}\right) - \left(\frac{2^4}{4} - 2\frac{2^3}{3} + \mu^2\frac{2^2}{2}\right)\right] \\ &\approx c \times 9.92 \approx 4 \end{aligned}$$

(c) It turns out that, $Var[X] = 4$. Suppose $X_1, \ldots, X_{81}$ are iid from $f_X(x)$ and let $\overline{X} = \frac{1}{81}\sum_{i=1}^{81} x_i$. What is the approximate distribution of $\overline{X}$? Give the name of the distribution and the value(s) of parameter(s). (5 points)

Answer: Normal distribution with mean 2.533 and variance 4/81.

(d) Approximate $P(\overline{X} < 2)$. (5 points)

Answer: The best answer is that this probability is zero since none of the X_i can be less than 2.

If you use the Central Limit Theorem, then you have

$$P\left(\overline{X} < 2\right) = P\left(Z < \frac{2 - 2.533}{2/9}\right) = P(Z < -2.4) \approx 0.0082.$$