Introduction to Mathematical Statistics I

Homework 3

Name : Freya Zou

3.14 The maximum patent life for a new drug is 17 years. Subtracting the length of time required by the FDA for testing and approval of the drug provides the actual patent life for the drug—that is, the length of time that the company has to recover research and development costs and to make a profit. The distribution of the lengths of actual patent lives for new drugs is given below:

Years, y	3	4	5	6	7	8	9	10	11	12	13
$p(y)$	.03	.05	.07	.10	.14	.20	.18	.12	.07	.03	.01

a Find the mean patent life for a new drug.

$$\mu = E(Y) = \sum_{\text{all } y} y\,p(y) = (3*0.03) + 4*0.05 + 5*.07 + 6*.10 + 7*0.14 + 8*0.2 + 9*0.18 + 10*.12 + 11*0.07 + 12*0.03 + 13*0.01$$
$$= 7.9$$

thus the mean patent life for a new drug is 7.90

b Find the standard deviation of $Y =$ the length of life of a randomly selected new drug.

$$Var(Y) = \sum_{\text{all } y} (y-\mu)^2 \cdot p(y)$$

$$= (3-7.9)^2 0.03 + (4-7.9)^2 0.05 + (5-7.9)^2 0.07 + (6-7.9)^2 * 0.10 + (7-7.9)^2 * .14 + (8-7.9)^2 * 0.20 + (9-7.9)^2 .18 + (10-7.9)^2 * 0.12 + (11-7.9)^2 .07 + (12-7.9)^2 * 0.03 + (13-7.9)^2 * 0.01$$

$$= 4.73$$

$$\sigma = \sqrt{var(Y)} = \sqrt{4.73} \approx 2.17$$

c What is the probability that the value of Y falls in the interval $\mu \pm 2\sigma$?

From part (a) and (b) we know $\mu = 7.9$ and $\sigma = 2.17$

First let calculate $\mu - 2\sigma = 7.9 - 2(2.17) = 7.90 - 4.34 = 3.56$

$\mu + 2\sigma = 7.9 + 2(2.17) = 7.9 + 4.34 = 12.24$

so the interval is approximately $[3.56, 12.24]$

we want to find the probability that the value falls in interval $[3.56, 12.24]$, we need to sum up

$$P(Y \text{ in } [3.56, 12.24]) = P(4) + P(5) + P(6) + P(7) + P(8) + P(9) + P(10) + P(11) + P(12)$$
$$= 0.05 + 0.07 + 0.10 + 0.14 + 0.2 + 0.18 + 0.12 + 0.07 + 0.03$$
$$= 0.96$$

Therefore, the probability that Y falls within 2σ is 96%

3.23 In a gambling game a person draws a single card from an ordinary 52-card playing deck. A person is paid \$15 for drawing a jack or a queen and \$5 for drawing a king or an ace. A person who draws any other card pays \$4. If a person plays this game, what is the expected gain?

- There are 4 Jacks and 4 queens, thus the probability is $P(\text{Jack or Queen}) = \frac{8}{52} = \frac{2}{13} \rightarrow$ Gain: 15\$
- There are 4 kings and 4 Aces, thus the probability is also $\frac{2}{13} \Rightarrow$ Gain 5\$
- Draw Any other card, there are $52-8-8=36$ other cards left, thus $P(\text{other cards}) = \frac{36}{52} = \frac{9}{13} \Rightarrow$ loss: \$4

$$E(X) = \sum_{\text{all } x} x P(x) = \left(15 \cdot \frac{12}{13}\right) + \left(5 \cdot \frac{12}{13}\right) + \left(-4 \cdot \frac{9}{13}\right) \approx 0.30769$$

On average, a person can expect to gain \$0.30769 in a long Run

3.30 Suppose that Y is a discrete random variable with mean μ and variance σ^2 and let $X = Y + 1$.

a Do you expect the mean of X to be larger than, smaller than, or equal to $\mu = E(Y)$? Why?

b Use Theorems 3.3 and 3.5 to express $E(X) = E(Y+1)$ in terms of $\mu = E(Y)$. Does this result agree with your answer to part (a)?

c Recalling that the variance is a measure of spread or dispersion, do you expect the variance of X to be larger than, smaller than, or equal to $\sigma^2 = V(Y)$? Why?

d Use Definition 3.5 and the result in part (b) to show that

$$V(X) = E\{[(X - E(X)]^2\} = E[(Y - \mu)^2] = \sigma^2;$$

that is, $X = Y + 1$ and Y have *equal* variances.

a. since $X = Y+1$, we expect the mean of X to be larger than the mean of Y, addiy a constant to a random variable shifts its mean by the contract

therefore: $E(X) = E(Y+1) = E(Y)+1 = \mu+1$

The result indicates that the mean of X is 1 unit larger than the mean of Y.

b. $E(X) = E(Y+1) = E(Y) + E(1)$, since $E(1) = 1$, we have $E(X) = E(Y)+1 = \mu+1$, This confirm that the mean of X is indeed 1 unit greater than the mean of Y, which agrees with our answer to part (a)

C. Variance meases the spread of a random variable, not its locaty Addy a constant shifts the distribution without affecty the spread. Therefe the variance of X should be equal to the variance of Y: $V(X) = V(Y) = \sigma^2$

d. $V(X) = E[(X - E(X))^2]$

substituting $X = Y+1$ and $E(X) = \mu+1$:

$$V(X) = E[((Y+1) - (\mu+1))^2]$$

$$= E[(Y-\mu)^2]$$

we know from the properties of variance that :

$$V(Y) = E[(Y-\mu)^2] = \sigma^2$$

thus $V(X) = \sigma^2$

This confirms that $X = Y+1$ and Y have equal variance.

3.40 The probability that a patient recovers from a stomach disease is .8. Suppose 20 people are known to have contracted this disease. What is the probability that

a exactly 14 recover?
b at least 10 recover?
c at least 14 but not more than 18 recover?
d at most 16 recover?

Given that $P(\text{recovers}) = 0.8$ $q(\text{not recovers}) = 0.2$

a. $P(X=x) = \binom{n}{x} p^x q^{(n-x)} = P(X=14) = \binom{20}{14} 0.8^{14} 0.2^{6} \approx 0.1090$

b. $P(X \geq 10) = 1 - P(X<10) = 1 - [P(X=0) + P(X=1) + P(X=2) \cdots + P(X=9)]$

$\Rightarrow 1 - [\binom{20}{0} 0.8^0 0.2^{20} + \binom{20}{1} 0.8^1 0.2^{19} + \binom{20}{2} 0.8^2 0.2^{18} + \cdots \binom{20}{9} 0.8^9 0.2^{11}]$.

$\Rightarrow \approx 0.9994$

c. we need: $P(14 \leq X < 18)$.

$P(X=x) = \binom{n}{x} p^x q^{(n-x)} = P(X=14) = \binom{20}{14} 0.8^{14} 0.2^{6} +$

$P(X=x) = \binom{n}{x} p^x q^{(n-x)} = P(X=15) = \binom{20}{15} 0.8^{15} 0.2^{5} +$

$P(X=x) = \binom{n}{x} p^x q^{(n-x)} = P(X=16) = \binom{20}{16} 0.8^{16} 0.2^{4} +$

$P(X=x) = \binom{n}{x} p^x q^{(n-x)} = P(X=17) = \binom{20}{17} 0.8^{17} 0.2^{3} +$

P(at least 14 not more than 18) $= P(X=14) + P(X=15) + P(X=16) + P(X=17)$

≈ 0.8441

d. we need to find $P(X \leq 16) = \sum_{x=0}^{16} \binom{n}{x_i} p^{x_i} q^{(n-x_i)}$

$\Rightarrow \binom{20}{0} 0.8^0 0.2^{20} + \binom{20}{1} 0.8^1 0.2^{19} + \binom{20}{2} 0.8^2 0.2^{18} + \cdots + \binom{20}{16} 0.8^{16} 0.2^{4}$

≈ 0.5886

3.70 An oil prospector will drill a succession of holes in a given area to find a productive well. The probability that he is successful on a given trial is .2.

a What is the probability that the third hole drilled is the first to yield a productive well?

b If the prospector can afford to drill at most ten wells, what is the probability that he will fail to find a productive well?

a. $P(X=3) = q^{x-1}p = 0.8^{3-1}\,0.2 = 0.128$

b. we have to find $P(X \le 10)$, we can sum the prob of sucesses on each of 10 trials as follow:

$P(\text{fail}) = 1-(q^{x-1}p) = 1-[P(X=1)+P(X=2)+\cdots+P(X=10)]$

$= 1-(0.8^0 0.2 + 0.8^1 0.2 + \cdots + 0.8^9 0.2)$

$\approx 1-0.8926 \approx 0.1074$

***3.85** Find $E[Y(Y-1)]$ for a geometric random variable Y by finding $d^2/dq^2\left(\sum_{y=1}^{\infty} q^y\right)$. Use this result to find the variance of Y.

For a Geometric random variable Y, with parameter p,

PGF: $S(q) = \sum_{y=1}^{\infty} q^y = \frac{q}{1-q}$, for $|q|<1$, to find $E[Y(Y-1)]$, we'll have to differentiate $S(q)$ twice

$S'(q) = \frac{d}{dq}\left(\frac{q}{1-q}\right) = \frac{1}{(1-q)^2}$

$\Rightarrow S''(q) = \frac{d}{dq}\left(\frac{1}{(1-q)^2}\right) = \frac{1}{(1-q)^3}$

Expectation $E(Y(Y-1))$, we know PGF is

$G_X(q) = E[q^Y] = \frac{pq}{1-(1-p)q}$, $|q| < \frac{1}{1-p}$

Applying the dervatives we found

$G''_Y(q) = \frac{2(1-p)^2}{[1-(1-p)q]^3}$, Plugg $q=1$

thus $E[Y(Y-1)] = \frac{2(1-p)^2}{p^2}$

Step 4 : variance Y we can recall $Var(Y) = E[Y^2] - (E[Y])^2$

1. $E[Y] = \frac{1}{p}$, 2. $E[Y^2] = E[Y(Y-1)] + E[Y] = \frac{2(1-p)^2}{p^2} + \frac{1}{p}$

we can simplyfy : $E[Y^2] = \frac{2(1-p)^2}{p^2} + \frac{1}{p}$

thus $Var(Y) = E[Y^2] - \left(\frac{1}{p}\right)^2 = \frac{2(1-p)^2}{p^2} + \frac{1}{p} - \frac{1}{p^2}$

$\Rightarrow Var(Y) = \frac{2(1-p)^2 + p - 1}{p^2}$

$\Rightarrow Var(Y) = \frac{2-4p+2p^2+p-1}{p^2}$

$\Rightarrow Var(Y) = \frac{2p^2-3p+1}{p^2}$

using the fact that $Var(Y)$ for a geometric distriency is $\frac{1-p}{p^2}$, this comfims that my result is consistent with the know result

3.167 Let Y be a random variable with mean 11 and variance 9. Using Tchebysheff's theorem, find

a a lower bound for $P(6 < Y < 16)$.

b the value of C such that $P(|Y - 11| \ge C) \le .09$.

Given $Var(X) = \sigma^2 = 9$, $E(X) = 11 = \mu$, $\sigma = 3$

a. $P(|Y-\mu| < k\sigma) \ge 1 - \frac{1}{k^2}$, we need to calculate distance the mean and bounds

$|Y-11| < 5$, Because 6 and 16 are both 5 units from 16

thus $k = \frac{5}{\sigma} = \frac{5}{3}$ we can apply Tchebysheff's theorem

$P(|Y-\mu| < k\sigma) \ge 1 - \frac{1}{(\frac{5}{3})^2} = 0.64$

b. $P(|Y-11| \ge c) \le \frac{\sigma^2}{C^2}$, we want $P(|Y-11| \ge c) \le 0.09$

thus $\frac{\sigma^2}{C^2} \le 0.09$, Given $\sigma^2 = 9$, we can substitute to the equation $\frac{9}{C^2} \le 0.09 \Rightarrow C = 10$, thus $C = 10$ such that $P(|Y-11| \ge c) \le 0.09$

3.177 For a certain section of a pine forest, the number of diseased trees per acre, Y, has a Poisson distribution with mean $\lambda = 10$. The diseased trees are sprayed with an insecticide at a cost of \$3 per tree, plus a fixed overhead cost for equipment rental of \$50. Letting C denote the total spraying cost for a randomly selected acre, find the expected value and standard deviation for C. Within what interval would you expect C to lie with probability at least .75?

(a) expected value for c

The total cost c can be expressed: $C = 50 + 3y$

since Y follows a Poisson distribution with $\lambda = 10$

$E(Y) = \lambda = 10$, now, the expected value of c is

$$E(c) = E(50+3Y) = 50 + 3E(Y)$$
$$= 50 + 3(10) = 80$$

so the expected value of c is 80\$.

(b) standard Deviation for c

$$Var(c) = Var(50+3Y) = 3^2 \cdot Var(Y)$$
$$\Rightarrow Var(C) = 9 \cdot 10 = 90$$
$$\sigma \text{ for } c = \sqrt{90} \approx 9.49$$

(c) within that interval would you expect C to lie with probability at least 0.75.

we can use Tscheff's inequality $P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$

we can subsitude: $P(|C-80| < k \cdot 9.49) \geq 0.75$

Applyiy chebyshe've inequality: $\frac{1}{k^2} \leq 0.25$

$$k^2 \geq 4$$
$$k \geq 2$$

Now, the interval is $80 \pm 2 \cdot 9.49 \Rightarrow 80 \pm 18.98$

thus we know the interval is $(61.02, 98.98)$

Therefore, the interval within wich c is expected to lie at least 0.75 probability is approximatey $(61.02, 98.98)$