Recitation 12

Random Variables

A random variable depends on the outcome of an uncertain event and takes specific values with specific probabilities (called a distribution)

Example: Let X denote the outcome of the roll of a fair die. Then X is a random variable with values {1, 2, 3, 4, 5, 6} each with a probability of 1/6.

Note: "Random" does not necessarily mean random with equal chance, we can have X be {1, 2, 3, 4, 5, 6} with probabilities {½, 1/10, 1/10, 1/10, 1/10, 1/10} (This can be written as $P(X=1) = ½$, $P(X=2) = 1/10$, etc. Where P is the Probability Mass Function)

Expectation

The Expected Value of a RV X, denoted as E[X] is the "average" value attained by the RV if we repeat X a large number of times. We denote the "average" or mean of X as μ_x

$$E[X] = \sum_{\text{all possible } x} xP(X = x).$$

Example: The expected value of the roll of a die is

$$1(\frac{1}{6}) + 2(\frac{1}{6}) + 3(\frac{1}{6}) + \cdots + 6(\frac{1}{6}) = 21/6 = 3.5.$$

Expectation

We also define the expected value for a function of a random variable. If g is a function (for example, $g(x) = x^2$), then the expected value of $g(X)$ is

$$E[g(X)] = \sum_{\text{all possible } x} g(x)P(X = x).$$

For example,

$$E[X^2] = \sum_{\text{all possible } x} x^2 P(X = x).$$

In general, $E[g(X)]$ is not the same as $g(E[X])$. In particular, $E[X^2]$ is not the same as $(E[X])^2$.

Ex: we roll a four-sided die, X = {1, 2, 3, 4}, P = {¼, ¼, ¼, ¼}

$E[X]^2 = [(1*(¼)) + (2*(¼)) + (3*(¼)) + (4*(¼))]^2$
$E[X^2] = [(1^2*(¼)) + (2^2*(¼)) + (3^2*(¼)) + (4^2*(¼))]$

Theorem 2.1 [Linearity of Expectations]: *For any finite collection of discrete random variables $X_1, X_2, \ldots, X_n$ with finite expectations,*

$$\mathbf{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbf{E}[X_i].$$

Proof: We prove the statement for two random variables X and Y; the general case follows by induction. The summations that follow are understood to be over the ranges of the corresponding random variables:

$$\begin{aligned}
\mathbf{E}[X+Y] &= \sum_i \sum_j (i+j)\Pr((X=i)\cap(Y=j)) \\
&= \sum_i \sum_j i\Pr((X=i)\cap(Y=j)) + \sum_i \sum_j j\Pr((X=i)\cap(Y=j)) \\
&= \sum_i i \sum_j \Pr((X=i)\cap(Y=j)) + \sum_j j \sum_i \Pr((X=i)\cap(Y=j)) \\
&= \sum_i i\Pr(X=i) + \sum_j j\Pr(Y=j) \\
&= \mathbf{E}[X] + \mathbf{E}[Y].
\end{aligned}$$

Variance

The **variance** of a random variable X is denoted by either $Var[X]$ or σ_X^2. (σ is the Greek letter sigma.) The variance is defined by

$$\sigma_X^2 = E[(X - \mu_X)^2];$$

this is the expected value of the squared difference between X and its mean. For a discrete distribution, we can write the variance as

$$\sigma_X^2 = \sum_x (x - \mu_X)^2 P(X = x).$$

An alternative expression for the variance (valid for both discrete and continuous random variables) is

$$\sigma_X^2 = E[(X^2)] - (\mu_X)^2.$$

This is the difference between the expected value of X^2 and the square of the mean of X.

The **standard deviation** of a random variable is the square-root of its variance and is denoted by σ_X. Generally speaking, the greater the standard deviation, the more spread-out the possible values of the random variable.

Problem 1

Let X be a discrete random variable with the following PMF

$$P_X(k) = \begin{cases} \frac{1}{4} & \text{for } k = -2 \\ \frac{1}{8} & \text{for } k = -1 \\ \frac{1}{8} & \text{for } k = 0 \\ \frac{1}{4} & \text{for } k = 1 \\ \frac{1}{4} & \text{for } k = 2 \\ 0 & \text{otherwise} \end{cases}$$

I define a new random variable Y as $Y = (X+1)^2$.

a. Find the range of Y.
b. Find the PMF of Y.

Problem 1

Let X be a discrete random variable with the following PMF

$$P_X(k) = \begin{cases} \frac{1}{4} & \text{for } k = -2 \\ \frac{1}{8} & \text{for } k = -1 \\ \frac{1}{8} & \text{for } k = 0 \\ \frac{1}{4} & \text{for } k = 1 \\ \frac{1}{4} & \text{for } k = 2 \\ 0 & \text{otherwise} \end{cases}$$

I define a new random variable Y as $Y = (X+1)^2$.

a. Find the range of Y.

b. Find the PMF of Y.

Here, the random variable Y is a function of the random variable X. This means that we perform the random experiment and obtain $X = x$, and then the value of Y is determined as $Y = (x+1)^2$. Since X is a random variable, Y is also a random variable.

a. To find R_Y, we note that $R_X = \{-2, -1, 0, 1, 2\}$, and

$$\begin{aligned} R_Y &= \{y = (x+1)^2 | x \in R_X\} \\ &= \{0, 1, 4, 9\}. \end{aligned}$$

b. Now that we have found $R_Y = \{0, 1, 4, 9\}$, to find the PMF of Y we need to find $P_Y(0), P_Y(1), P_Y(4)$, and $P_Y(9)$:

$$\begin{aligned} P_Y(0) &= P(Y = 0) = P((X+1)^2 = 0) \\ &= P(X = -1) = \tfrac{1}{8}; \\ P_Y(1) &= P(Y = 1) = P((X+1)^2 = 1) \\ &= P\big((X = -2) \text{ or } (X = 0)\big); \\ &P_X(-2) + P_X(0) = \tfrac{1}{4} + \tfrac{1}{8} = \tfrac{3}{8}; \\ P_Y(4) &= P(Y = 4) = P((X+1)^2 = 4) \\ &= P(X = 1) = \tfrac{1}{4}; \\ P_Y(9) &= P(Y = 9) = P((X+1)^2 = 9) \\ &= P(X = 2) = \tfrac{1}{4}. \end{aligned}$$

Again, it is always a good idea to check that $\sum_{y \in R_Y} P_Y(y) = 1$. We have

$$\sum_{y \in R_Y} P_Y(y) = \frac{1}{8} + \frac{3}{8} + \frac{1}{4} + \frac{1}{4} = 1.$$

Problem 2

Let X be a discrete random variable with the following PMF

$$P_X(k) = \begin{cases} 0.1 & \text{for } k = 0 \\ 0.4 & \text{for } k = 1 \\ 0.3 & \text{for } k = 2 \\ 0.2 & \text{for } k = 3 \\ 0 & \text{otherwise} \end{cases}$$

a. Find EX.

b. Find Var(X).

c. If $Y = (X - 2)^2$, find EY.

Problem 2a

Let X be a discrete random variable with the following PMF

$$P_X(k) = \begin{cases} 0.1 & \text{for } k = 0 \\ 0.4 & \text{for } k = 1 \\ 0.3 & \text{for } k = 2 \\ 0.2 & \text{for } k = 3 \\ 0 & \text{otherwise} \end{cases}$$

a. Find EX.
b. Find Var(X).
c. If $Y = (X - 2)^2$, find EY.

$$\begin{aligned} EX &= \Sigma_{x_k \in R_X} x_k P_X(x_k) \\ &= 0(0.1) + 1(0.4) + 2(0.3) + 3(0.2) \\ &= 1.6 \end{aligned}$$

Problem 2b

Let X be a discrete random variable with the following PMF

$$P_X(k) = \begin{cases} 0.1 & \text{for } k = 0 \\ 0.4 & \text{for } k = 1 \\ 0.3 & \text{for } k = 2 \\ 0.2 & \text{for } k = 3 \\ 0 & \text{otherwise} \end{cases}$$

a. Find EX.
b. Find Var(X).
c. If $Y = (X-2)^2$, find EY.

We can use Var(X) $= EX^2 - (EX)^2 = EX^2 - (1.6)^2$. Thus we need to find EX^2. Using LOTUS, we have

$$EX^2 = 0^2(0.1) + 1^2(0.4) + 2^2(0.3) + 3^2(0.2) = 3.4$$

Thus, we have

$$\mathrm{Var}(X) = (3.4) - (1.6)^2 = 0.84$$

Problem 2c

Let X be a discrete random variable with the following PMF

$$P_X(k) = \begin{cases} 0.1 & \text{for } k = 0 \\ 0.4 & \text{for } k = 1 \\ 0.3 & \text{for } k = 2 \\ 0.2 & \text{for } k = 3 \\ 0 & \text{otherwise} \end{cases}$$

a. Find EX.
b. Find Var(X).
c. If $Y = (X - 2)^2$, find EY.

$Y = (X-2)*(X-2) = X^2 - 4X + 4$

$E[Y] = E[X^2 - 4X + 4] = E[X^2] - E[X] + E[4] = E[X^2] - E[X] + 4$ using linearity of expectation

$E[X] = 0*(0.1) + 1*(0.4) + 2*(0.3) + 3*(0.2) = (0.4) + 2*(0.3) + 3*(0.2)$

$E[X^2] = (0^2)*(0.1) + (1^2)*(0.4) + (2^2)*(0.3) + (3^2)*(0.2) = (0.4) + 4*(0.3) + 9*(0.2)$

$E[Y] = [(0.4) + 2*(0.3) + 3*(0.2)] + [(0.4) + 4*(0.3) + 9*(0.2)] + 4$

Challenge Problem

You take an exam that contains 20 multiple-choice questions. Each question has 4 possible options. You know the answer to 10 questions, but you have no idea about the other 10 questions so you choose answers randomly. Your score X on the exam is the total number of correct answers. Find the PMF of X. What is $P(X > 15)$?

Challenge Problem

You take an exam that contains 20 multiple-choice questions. Each question has 4 possible options. You know the answer to 10 questions, but you have no idea about the other 10 questions so you choose answers randomly. Your score X on the exam is the total number of correct answers. Find the PMF of X. What is $P(X > 15)$?

Let's define the random variable Y as the number of your correct answers to the 10 questions you answer randomly. Then your total score will be $X = Y + 10$. First, let's find the PMF of Y. For each question your success probability is $\frac{1}{4}$. Hence, you perform 10 independent $\text{Bernoulli}(\frac{1}{4})$ trials and Y is the number of successes. Thus, we conclude $Y \sim \text{Binomial}(10, \frac{1}{4})$, so

$$P_Y(y) = \begin{cases} \binom{10}{y}(\frac{1}{4})^y(\frac{3}{4})^{10-y} & \text{for } y = 0, 1, 2, 3, \ldots, 10 \\ 0 & \text{otherwise} \end{cases}$$

Now we need to find the PMF of $X = Y + 10$. First note that $R_X = \{10, 11, 12, \ldots, 20\}$. We can write

$$\begin{aligned} P_X(10) &= P(X = 10) = P(Y + 10 = 10) \\ &= P(Y = 0) = \binom{10}{0}\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^{10-0} = \left(\frac{3}{4}\right)^{10}; \\ P_X(11) &= P(X = 11) = P(Y + 10 = 11) \\ &= P(Y = 1) = \binom{10}{1}\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^{10-1} = 10\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^9. \end{aligned}$$

So, you get the idea. In general for $k \in R_X = \{10, 11, 12, \ldots, 20\}$,

$$\begin{aligned} P_X(k) &= P(X = k) = P(Y + 10 = k) \\ &= P(Y = k - 10) = \binom{10}{k-10}\left(\frac{1}{4}\right)^{k-10}\left(\frac{3}{4}\right)^{20-k}. \end{aligned}$$

Challenge Problem

You take an exam that contains 20 multiple-choice questions. Each question has 4 possible options. You know the answer to 10 questions, but you have no idea about the other 10 questions so you choose answers randomly. Your score X on the exam is the total number of correct answers. Find the PMF of X. What is $P(X > 15)$?

To summarize,

$$P_X(k) = \begin{cases} \binom{10}{k-10}(\frac{1}{4})^{k-10}(\frac{3}{4})^{20-k} & \text{for } k = 10, 11, 12, \ldots, 20 \\ 0 & \text{otherwise} \end{cases}$$

In order to calculate $P(X > 15)$, we know we should consider $y = 6, 7, 8, 9, 10$

$$P_Y(y) = \begin{cases} \binom{10}{y}(\frac{1}{4})^{y}(\frac{3}{4})^{10-y} & \text{for } y = 6, 7, 8, 9, 10 \\ 0 & \text{otherwise} \end{cases}$$

$$P_X(k) = \begin{cases} \binom{10}{k-10}(\frac{1}{4})^{k-10}(\frac{3}{4})^{20-k} & \text{for } k = 16, 17, \ldots, 20 \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} P(X > 15) &= P_X(16) + P_X(17) + P_X(18) + P_X(19) + P_X(20) \\ &= \binom{10}{6}(\frac{1}{4})^6(\frac{3}{4})^4 + \binom{10}{7}(\frac{1}{4})^7(\frac{3}{4})^3 + \binom{10}{8}(\frac{1}{4})^8(\frac{3}{4})^2 \\ &+ \binom{10}{9}(\frac{1}{4})^9(\frac{3}{4})^1 + \binom{10}{10}(\frac{1}{4})^{10}(\frac{3}{4})^0. \end{aligned}$$

FINAL QUIZ!!

- You will have 25 minutes to complete the Quiz

- Final Exam will be on Dec 20th