STAT230: Principles of Probability

Unit 5: Discrete Random Variables

Department of Statistics

United Arab Emirates University

March 14, 2022

Outline

- Random Variables RV
- Discrete Random Variables DRV
- Expected Value and Variance $E(X)$ and VAR(X)
- VAR(X)
- Binomial Distribution $\mathcal{B}(n, p)$
- Poisson Distribution $\mathcal{P}(\lambda)$
- Negative Binomial Distribution $\mathcal{B}^{-}(r, n, p)$
- Geometric Distribution $\mathcal{G}(p)$
- Simulations

Random Variables

- Frequently, when an experiment is performed, we are interested mainly in some function of the outcome as opposed to the actual outcome itself.
- For instance, in tossing dice, we are often interested in the sum of the two dice and are not really concerned about the separate values of each die.
- These quantities of interest, or, more formally, these real-valued functions defined on the sample space, are known as *random variables*.
- Example 1: Suppose that our experiment consists of tossing 3 fair coins. If we let Y denote the number of heads that appear, then Y is a random variable taking on one of the values 0, 1, 2, and 3 with respective probabilities $p_0 = P(Y = 0) = ...$, $p_1 = P(Y = 1) = ...$, $p_2 = P(Y = 2) = ...$, $p_3 = P(Y = 3) = ...$
- Example 2: Toss a Coin Experiment OBJ

- A random variable that can take on at most a *countable* number of possible values is said to be *discrete*.
- For a discrete random variable (DRV) X, we define the probability mass function (pmf) $p(a)$ of X by

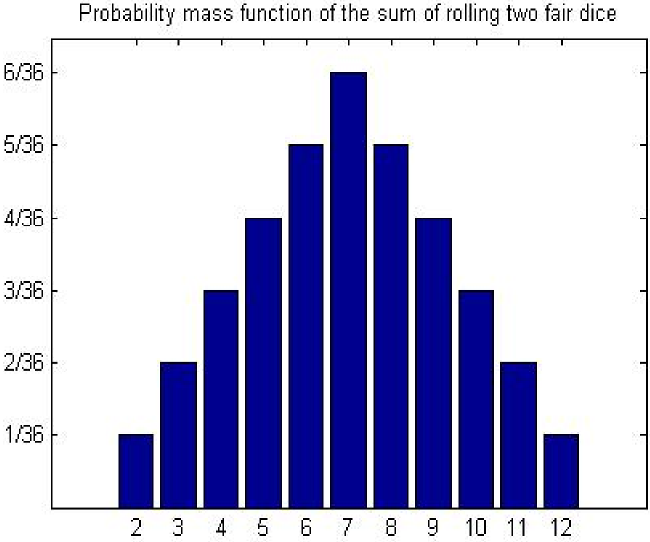
$$p(a) \quad = \quad \mathcal{P}\{X = a\} \tag{1}$$

- Example 2: The pmf of the random variable representing the sum when two dice are rolled can be represented in two ways.
- First, as a table:

x	2	3	4	5	6	7	8	9	10	11	12
$p(x)$	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/35$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$

Discrete Random Variables

- Second, as a graph:



Exercise 1

- A system consists of 2 components connected in parallel, then at least one must work correctly for the system to work correctly. Each component operates correctly with probability 0.8 and independent of the other. Let X be the number of components that work correctly. Find the probability distribution of X.
- Solution: X can take on only three possible values; 0, 1, or 2. Let E_i denote the event that component i works correctly. Then $P(E_i) = 0.8$. Thus, we have

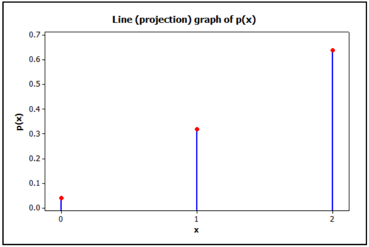
$$\begin{aligned}
p(0) &= P(E_1' \cap E_2') = P(E_1')P(E_2') \\
&= (0.2)(0.2) = 0.04 \\
p(1) &= P(E_1' \cap E_2) + P(E_1 \cap E_2') = P(E_1')P(E_2) + P(E_1)P(E_2') \\
&= (0.2)(0.8) + (0.8)(0.2) = 0.32 \\
p(2) &= P(E_1 \cap E_2) = P(E_1)P(E_2) \\
&= (0.8)(0.8) = 0.64
\end{aligned}$$

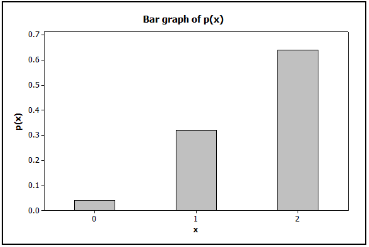
Exercise 1 (Cont.)

So the probability mass function of X is given by

X	0	1	2
$P(X)$	0.04	0.32	0.64

The figure below shows the pictorial presentation of the probability mass function of X.





Discrete Random Variables: CDF

Definition (Cumulative Distribution Function)

The *cumulative distribution function* (cdf), denoted by $\mathcal{F}$, can be expressed in terms of $p(a)$ by

$$\mathcal{F}(a) = \sum_{x \leq a} p(x) = \sum_{x=-\infty}^{a} p(x) \tag{2}$$

- Example 3: If X has a pmf given by $P(1) = 1/4$, $p(2) = 1/2$, $p(3) = p(4) = 1/8$. Find the cdf of X.
- Solution: The cdf of X is given by:

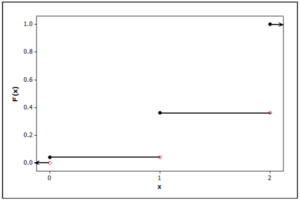
$$\mathcal{F}(a) = \begin{cases} 0 & \text{if } a < 1; \\ \frac{1}{4} & \text{if } 1 \leq a < 2; \\ \frac{3}{4} & \text{if } 2 \leq a < 3; \\ \frac{7}{8} & \text{if } 3 \leq a < 4; \\ 1 & \text{if } a \geq 4. \end{cases}$$

Exercise 2:

- Find and plot the CDF of the RV X defined in Exercise 1.
- Solution: The cdf of X in Exercise 1 is given by

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.04 & 0 \leq x < 1 \\ 0.36 & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

- The graph of the CDF is given by



- Question: Plot the CDF of Example 3. OBJ

The Expected Value

Definition (The Expected Value (Mean))

If X is a random variable with pmf $p(X)$, then the expected value (the mean) of X denoted by $E(X)$ or μ is given by

$$E(X) = \mu = \sum_x xp(x)$$

The Expected Value of a Function

- Let $h(x)$ be any function and let Y denote the random variable $Y = h(X)$, then the expected value of $Y = h(X)$ is computed by

$$E(Y) = \sum_x h(x)p(x)$$

- If the function $h(x)$ is linear, i.e. $h(x) = ax + b$, where a is a multiplicative constant and b is an additive constant, then

$$E(Y) = E(aX + b) = aE(X) + b$$

Variance and Standard Deviation

Variance and Standard Deviation

- The variance of X, denoted by Var(X) or σ^2 is given by

$$\begin{aligned} \text{Var}(X) = \sigma^2 &= E(X-\mu)^2 = \sum_x (x-\mu)^2 p(x) \\ &= E(X^2) - \mu^2 \end{aligned}$$

- The standard deviation of X is given by $\sigma = \sqrt{\sigma^2}$.
- Let $h(x) = ax + b$, then

$$\text{Var}[h(X)] = \text{Var}(aX+b) = a^2\text{Var}(X)$$

Note: For a random variable X having a pmf $p(x)$, the mean μ is a measure of the center of the pmf, and the variance σ^2 is a measure of the dispersion, or variability in the distribution. Note that these two measures do not uniquely identify a pmf. That is, two different pmfs can have the same mean and variance.

Examples: $E(X)$ and Var(X)

The probability distribution of the number of daily network blackouts is given by

X	0	1	2
$P(X)$	0.7	0.2	0.1

(a) Find the expected value and variance of the number of network blackouts.

(b) Solution: The expected value of X is

$$E(X) = \mu = \sum xp(x) = (0)(0.7) + (1)(0.2) + (2)(0.1) = 0.4$$

To compute the variance we need to find $E(X^2)$ first

$$E(X^2) = \sum x^2 p(x) = (0^2)(0.7) + (1^2)(0.2) + (2^2)(0.1) = 0.6$$

Hence, the variance of X is

$$\text{Var}(X) = E(X^2) - \mu^2 = 0.6 - 0.4^2 = 0.44$$

Examples: $E(X)$ and Var(X)

(b) A small internet trading company estimates that each network blackout results in a \$500 loss. Compute expectation and variance of this company's daily loss due to blackouts.

The daily loss due to blackouts is given by $Y = h(X) = 500X$. One could find the pmf of Y and use it to compute the expected value and variance of Y. Based on the pmf of X, the pmf of Y will be

Y	0	500	1000
$P(Y)$	0.7	0.2	0.1

However, using properties of expectation and variance, $E(Y)$ and Var(Y) can be computed without first computing the pmf of Y as follows:

$$E(Y) = E(500X) = 500E(X) = (500)(0.4) = 200$$

$$\text{Var}(Y) = \text{Var}(500X) = 500^2\text{Var}(X) = (500^2)(0.44) = 110,000$$

Exercises: $E(X)$ and $\mathrm{Var}(X)$

(a) Find $E(X)$ and $\mathrm{Var}(X)$, where X is the outcome when we roll a fair die.

(b) A school class of 120 students is driven in 3 buses to a symphonic performance. There are 36 students in one of the buses, 40 in another, and 44 in the third bus. When the buses arrive, one of the 120 students is randomly chosen. Let X denote the number of students on the bus of that randomly chosen student, and find $E(X)$ and σ_X.

(c) We say that I is an indicator variable for the event A if

$$I = \begin{cases} 0 & \text{if } A \text{ occurs} \\ 1 & \text{if } A^c \text{ occurs} \end{cases}$$

Find the $E(X)$ and $\mathrm{Var}(X)$.

OBJ

Binomial Distribution

- A Bernoulli experiment is a random experiment, the outcome of which can be classified in one of two mutually exclusive and exhaustive ways, say, "1=success" or "0=failure." Let Y be the number of success on a Bernoulli trial, then Y is called the Bernoulli random variable.

- If a sequence of n independent Bernoulli trials is performed under the same condition, we call a set of n Bernoulli trials a Binomial experiment.

Binomial Distribution

Binomial Experiment

An experiment is called a Binomial experiment if it satisfies the following 4 conditions:

1. The experiment consists of n Bernoulli trials.
2. Each trial results in a success (S) or a failure (F).
3. The trials are independent.
4. The probability of a success, p, is fixed throughout n trials.

Binomial Distribution $Bin(n, p)$

- Given a Binomial experiment consisting of n Bernoulli trials with success probability p, the Binomial random variable X associated with this experiment is defined as the number of successes among the n trials.

- The random variable X has the Binomial Distribution with parameters n and p; denoted by $X \sim Bin(n, p)$.

- The behavior of Binomial Distribution with different n and p

Binomial Distribution

- The probability mass function of $Bin(n,p)$ is given by

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0, 1, \ldots, n \text{ and } 0 < p < 1$$

- If $X \sim Bin(n,p)$, then:

$$\begin{aligned} E(X) &= np \\ \mathrm{Var}(X) &= np(1-p) \end{aligned}$$

Binomial Distribution: Examples

1. Five fair coins are flipped. If the outcomes are assumed independent.
   (1) Find the probability mass function of the number of heads obtained.
   (2) Find the probability that at least 3 heads are obtained. .
   (3) Use Binomial Calculator to find the probability that at most 2 heads are obtained.
2. Solt.:
   (1) Let $X =$ The number of heads in 5 tossed coins. $X \sim Bin(5, 0.5)$.
   $P(X=0) = .5^5 = 0.0313$, $P(X=1) = \binom{5}{1}.5^5 = 0.1563$,
   $P(X=2) = \binom{5}{2}.5^5 = 0.3125$, $P(X=3) = \binom{5}{3}.5^5 = 0.3125$,
   $P(X=4) = \binom{5}{4}.5^5 = 0.1563$, $P(X=5) = \binom{5}{5}.5^5 = 0.0313$.
   (2) $P(X \geq 3) = P(X=3) + P(X=4) + P(X=5) = 0.5$.
   (3) Using the statistical calculator
   $P(X \leq 2) = P(X=0) + P(X=1) + P(X=2) = 0.1875$

Binomial Distribution: Examples

1. It is known that screws produced by a certain company will be defective with probability .01, independently of each other. The company sells the screws in packages of 10 and offers a money-back guarantee that at most 1 of the 10 screws is defective. What proportion of packages sold must the company replace? Use the Binomial Calculator or Statistical Tables.
2. Multiple-Choice Test.
3. Juvenile Delinquents.
4. The following gambling game, known as the wheel of fortune (or chuck-a-luck), is quite popular at many carnivals and gambling casinos: A player bets on one of the numbers 1 through 6. Three dice are then rolled, and if the number bet by the player appears i times, $i = 1, 2, 3$, then the player wins i units; if the number bet by the player does not appear on any of the dice, then the player loses 1 unit. Is this game fair to the player?

Poisson Distribution

- The Poisson distribution models the number of occurrences of an event when there is a known average rate per unit time or space λ.

- The requirements for a Poisson distribution are that:
  (a) no two events can occur simultaneously,
  (b) events occur independently in different intervals, and
  (c) the expected number of events in each time interval remain constant.

Poisson Distribution

Poisson Distribution

- The probability mass function of $Poisson(\lambda)$ is given by

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \qquad x = 0, 1, \ldots$$

- If $X \sim Poisson(\lambda)$, then:

$$\begin{aligned} E(X) &= \lambda \\ V(X) &= \lambda \end{aligned}$$

Poisson Distribution: Examples

1. The number of customers arriving at a service counter within one-hour period.
2. The number of typographical errors in a book counted per page.
3. The number of email messages received at the technical support center daily.
4. The number of traffic accidents that occur on Al Ain-Dubai road during a month.
5. Poisson Distribution Simulator

Example

Messages arrive at an electronic message center at random times, with an average of 9 messages per hour.

(a) What is the probability of receiving exactly five messages during the next hour?
Solt.: The number of messages received in an hour, X is modeled by Poisson distribution with $\lambda = 9$, i.e. $X \sim Poisson(9)$.

$$P(X = 5) = p(5) = \frac{9^5 e^{-9}}{5!} = 0.0607$$

Example

(b) What is the probability that more than 10 messages will be received within the next two hours?
Solt.: The number of messages received within a 2-hour period, Y is another Poisson distribution with $\lambda = (2)(9) = 18$, i.e. $Y \sim Poisson(18)$.

$$\begin{aligned} P(Y > 10) &= 1 - P(Y \leq 10) = 1 - F(10) \\ &= 1 - \sum_{k=0}^{10} \frac{18^k e^{-18}}{k!} \\ &= 1 - 0.0304 = 0.9696 \end{aligned}$$

(c) Use Statistical Tables to solve questions (a) and (b).

1. Develop a real life example in which you can easily apply:
   (a) Group 1: Poisson distribution.
   (b) Group 2: Binomial distribution.
   (c) Group 3: Poisson distribution
2. In each case, propose two problems which can be solved using the Statistical Calculator.
3. Can you propose an idea in which you can mix both distributions? (extra)

Geometric Distribution

- Suppose that independent trials, each having a probability p, $0 < p < 1$, of being a success, are performed until a success occurs.
- Example: The first head in tossing coin several times.
- Then, *Geometric distribution* models the number of trials performed until a success occurs.

Geometric Distribution

- The probability mass function of $Geometric(p)$ denoted also $\mathcal{G}(p()$ is given by

$$p(x) = (1-p)^{x-1}p, \qquad x = 1, \ldots$$

- If $X \sim \mathcal{G}(p)$, then:

$$\begin{aligned} E(X) &= \frac{1}{p} \\ V(X) &= \frac{q}{p^2} \end{aligned}$$

Geometric Distribution: Example

Suppose that the probability of engine malfunction during any one-hour period is $p = .02$. Find the probability that a given engine will survive two hours.

Solution: Letting Y denote the number of one-hour intervals until the first malfunction, we have

$$\text{P(survive two hours)} = P(Y \geq 3) = \sum_{y=3}^{\infty} p(y)$$

Then,

$$\begin{aligned}\text{P(survive two hours)} &= 1 - \sum_{y=1}^{2} p(y) = 1 - p - qp \\ &= 1 - 0.02 - (0.98)(0.02) = 0.9604.\end{aligned}$$

Now, find the mean and standard deviation of Y?

Geometric Distribution: Example

An urn contains N white and M black balls. Balls are randomly selected, one at a time, until a black one is obtained. If we assume that each ball selected is replaced before the next one is drawn, what is the probability that:

(a) exactly n draws are needed?

(b) at least k draws are needed?

Solution: If we let X denote the number of draws needed to select a black ball, then X is a $\mathcal{G}(p)$ with $p = M/(M+N)$. Hence,

(a) $P(X=n) = (1 - M/(M+N))^{n-1} M/(M+N) = \frac{MN^{n-1}}{(M+N)^n}$.

(b)

$$
\begin{aligned}
P(X \geq k) &= \frac{M}{M+N} \sum_{n=k}^{\infty} \left(\frac{N}{M+N}\right)^{(n-1)} \\
&= \left(\frac{M}{M+N}\right)\left(\frac{N}{M+N}\right)^{k-1} / \left(1 - \frac{N}{M+N}\right) = \left(\frac{N}{M+N}\right)^{(k-1)}.
\end{aligned}
$$

(c) Or simply, since the probability that at least k trials are necessary to obtain a success is equal to the probability that the first $k-1$ trials are all failures, then $P(X \geq k) = (1-p)^{k-1}$.

Negative Binomial Distribution

- Suppose that independent trials, each having probability p, $0 < p < 1$, of being a success are performed until a total of r successes is accumulated.
- Example: The third head in tossing coin several times.
- Then, *Negative Binomial distribution* models the number of trials performed until a the r^{th} success occurs.

Negative Binomial Distribution

- The probability mass function of Negative Binomial RV, denoted also $\mathcal{B}^{-}(r, p)$, is given by

$$p(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \qquad x = r, r+1, \ldots$$

- If $X \sim \mathcal{B}^{-}(r, p)$, then:

$$\begin{aligned} E(X) &= \frac{r}{p} \\ V(X) &= \frac{rq}{p^2} \end{aligned}$$

Negative Binomial Distribution: Example

A machine produces 1% defective parts. Using the statistical calculator, calculate the probability that

(a) 10 parts have to be selected until to get 2 defective parts.

(b) Between 20 to 25 parts have to be selected to get 2 defective parts.

Solution:

Simulations: Using R to compute probabilities

For all of the random variables that we have mentioned so far (and many more), R has built in capabilities of computing probabilities. The syntax is broken down into two pieces: the root and the prefix. The root determines which random variable that we are talking about, and here are the names of the ones that we have covered so far:

(a) binom is binomial
(b) geom is geometric
(c) pois is Poisson
(d) nbinom is Negative Binomial

The available prefixes are

(a) p computes the cumulative distribution
(b) d computes pdf or pmf
(c) r samples from the rv
(d) q quantile function

Simulations: Examples

Let X be binomial with $n = 10$ and $p = .3$.

- Compute $P(X = 5)$.
  We are interested in the pmf of a binomial RV, so we will use the R command dbinom as follows:
  $dbinom(5, 10, .3)$
- Compute $P(1 \leq X \leq 5)$.
  We want to sum
  $$\sum_{k=1}^{5} P(X = k).$$
  We could do that the long way as follows:
  $$dbinom(5, 10, .3) + dbinom(4, 10, .3) + \dots$$

Simulations: Examples

Note that dbinom also allows vectors as arguments, so we can compute all of the needed probabilities like this:

$$dbinom(1:5, 10, .3)$$

And, all we need to do is add those up:

$$sum(dbinom(1:5, 10, .3))$$

Check that we got the same answer.

- We can also use the command pbinom.