4. Random Variables

4. Random Variables

In many situations when an experiment is performed, we are interested in some function of the outcome rather than the outcome itself. Here are some examples:

Example When we roll a pair of dice, let's say we are not interested in the numbers that are obtained in each die but we are only interested in the sum of the numbers.

Example There are 20 questions in a multiple choice paper. Each question has 5 alternatives. A student answers all 20 questions by randomly and independently choosing one alternative out of 5 in each question. We are interested in $X :=$ number of correct answers.

4.1 Random Variables

Definitions *A **random variable**, X, is a mapping from the sample space to real numbers.*

We always use capital letters like X, Y, Z to denote random variables.

The range of a rv, denoted by χ, is the collection of all possible values it can take on. For instance,

- $X \to \{0, 1, \dots, n\}$; $Y \to \{1, 2, 3, \dots, \}$; $Z \to [0, \infty)$.
- **We then can use the range of rv to classify it to be a**
  - **Continuous rv, or**
  - **Discrete rv.**

DISCRETE RANDOM VARIABLE

- It is a rv that has a finite or countable range. → **Chapter 4**
  - The number of defective items, the number of sales for a store, ...

CONTINUOUS RANDOM VARIABLE

- It is a rv whose range is an interval over the real line. → **Chapter 5**
  - Weight of an item, time until failure of a mechanical component, length of an object, ...

4.1 Random Variables

Class Discussion An urn contains 20 chips numbered from 1 to 20. Three chips are chosen at random from this urn. Let X be the largest number among the three chips drawn.

What is the range of X?

Choice A	Choice B	Choice C
$[1, 20]$	$\{1, 2, 3, \cdots, 20\}$	$\{3, 4, 5, \cdots, 20\}$

4.1 Random Variables

Example Suppose we toss 3 fair coins. Let Y denote the number of heads appearing, then Y takes values 0, 1, 2, 3.
And

$$\begin{aligned}
P(Y=0) &= P((T,T,T)) = \frac{1}{8} \\
P(Y=1) &= P((H,T,T),(T,H,T),(T,T,H)) = \frac{3}{8} \\
P(Y=2) &= P((H,H,T),(H,T,H),(T,H,H)) = \frac{3}{8} \\
P(Y=3) &= P((H,H,H)) = \frac{1}{8}
\end{aligned}$$

4.1 Random Variables

Example An urn contains 20 chips numbered from 1 to 20. Three chips are chosen at random from this urn. Let X be the largest number among the three chips drawn. Then X takes values from $3, 4, \ldots, 20$.

Suppose a game is that, you will win if the largest number obtained is at least 17. What is the probability of winning?

4.1 Random Variables

Example Three balls are randomly chosen from an urn containing 3 white, 3 red, and 5 black balls. Suppose that we win \$1 for each white ball selected and lose \$1 for each red selected. If we let X denote our total winnings from the experiment, then X is a random variable taking on the possible values $0, \pm 1, \pm 2, \pm 3$ with respectively probabilities

One can easily check that the total probability is 1. Moreover, the probability that we win money is given by

4.2 Discrete Random Variables

Definitions *A random variable is said to be **discrete** if the range of X is either finite or countably infinite.*

We always use capital letters like X, Y, Z to denote random variables.

We always use lower-case letters like x, y, z to denote the possible values of the random variable.

Definitions *Suppose that random variable X is discrete, taking values $x_1, x_2, \ldots$, then the **probability mass function** of X, denoted by p_X (or simply as p if the context is clear), is defined as*

$$p_X(x) = \begin{cases} P(X = x) & \text{if } x = x_1, x_2, \ldots \\ 0 & \text{otherwise} \end{cases}.$$

When writing p.m.f., it is important to determine its range first.

(Properties of the probability mass function)

(i) $p_X(x_i) \geq 0$; for $i = 1, 2, \ldots$;

(ii) $p_X(x) = 0$; for other values of x;

But the domain of p.m.f. is the entire real line.

(iii) Since X must take on one of the values of x_i, $\sum_{i=1}^{\infty} p_X(x_i) = 1$.

4.2 Discrete Random Variables

Example Suppose a random variable X only takes values $0, 1, 2\ldots$. If the probability mass function of X (p.m.f. of X) is of the form:

$$p(k) = c\frac{\lambda^k}{k!}, \quad k = 0, 1, \ldots$$

where $\lambda > 0$ is a fixed positive value and c is a suitably chosen constant.

(a) What is this suitable constant? (b) Compute $P(X = 0)$ and (c) $P(X > 2)$.

4.2 Discrete Random Variables

Definition *The* ***cumulative distribution function*** *of X, abbreviated to* ***distribution function*** *(d.f.) of X, (denoted as F_X or F if context is clear) is defined as*

$$F_X : \mathbb{R} \longrightarrow \mathbb{R}$$

The range of c.d.f. is [0, 1]

where

$$F_X(x) = P(X \leq x) \quad \textit{for } x \in \mathbb{R}.$$

Remark Suppose that X is discrete and takes values $x_1, x_2, x_3, \ldots$ where $x_1 < x_2 < x_3 < \ldots$. Note then that F is a step function, that is, F is constant in the interval $[x_{i-1}, x_i)$ (F takes value $p(x_1) + \cdots + p(x_{i-1})$), and then take a jump of size $= p(x_i)$.

Important **To determine the c.d.f. F(x), it suffices to consider its values on the following intervals**

$$(-\infty, x_1), [x_1, x_2), [x_2, x_3), [x_3, x_4), \cdots, [x_n, \infty)$$

4.2 Discrete Random Variables

Given p.m.f., how to obtain c.d.f.?

Example if X has a probability mass function given by

p.m.f. $$p(1) = \frac{1}{4}, \quad p(2) = \frac{1}{2}, \quad p(3) = \frac{1}{8}, \quad p(4) = \frac{1}{8},$$

4.2 Discrete Random Variables

Given p.m.f., how to obtain c.d.f.?

Example if X has a probability mass function given by

p.m.f. $$p(1) = \frac{1}{4}, \quad p(2) = \frac{1}{2}, \quad p(3) = \frac{1}{8}, \quad p(4) = \frac{1}{8},$$

4.2 Discrete Random Variables

Given p.m.f., how to obtain c.d.f.?

Example *Let X be the sum of the numbers obtained by rolling 2 fair dice.*

p.m.f.

x	2	3	4	5	6	7	8	9	10	11	12
P(X=x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

In this case

$$x_1 = 2, x_2 = 3, x_3 = 4, x_4 = 5, x_5 = 6, x_6 = 7, x_7 = 8, x_8 = 9, x_9 = 10, x_{10} = 11, x_{11} = 12$$

Therefore, it suffices to consider the values of F(x) in the following intervals.

$$(-\infty, 2), [2, 3), [3, 4), [4, 5), [5, 6), [6, 7), [7, 8), [8, 9), [9, 10), [10, 11), [11, 12), [12, \infty)$$

4.2 Discrete Random Variables

Given p.m.f., how to obtain c.d.f.?

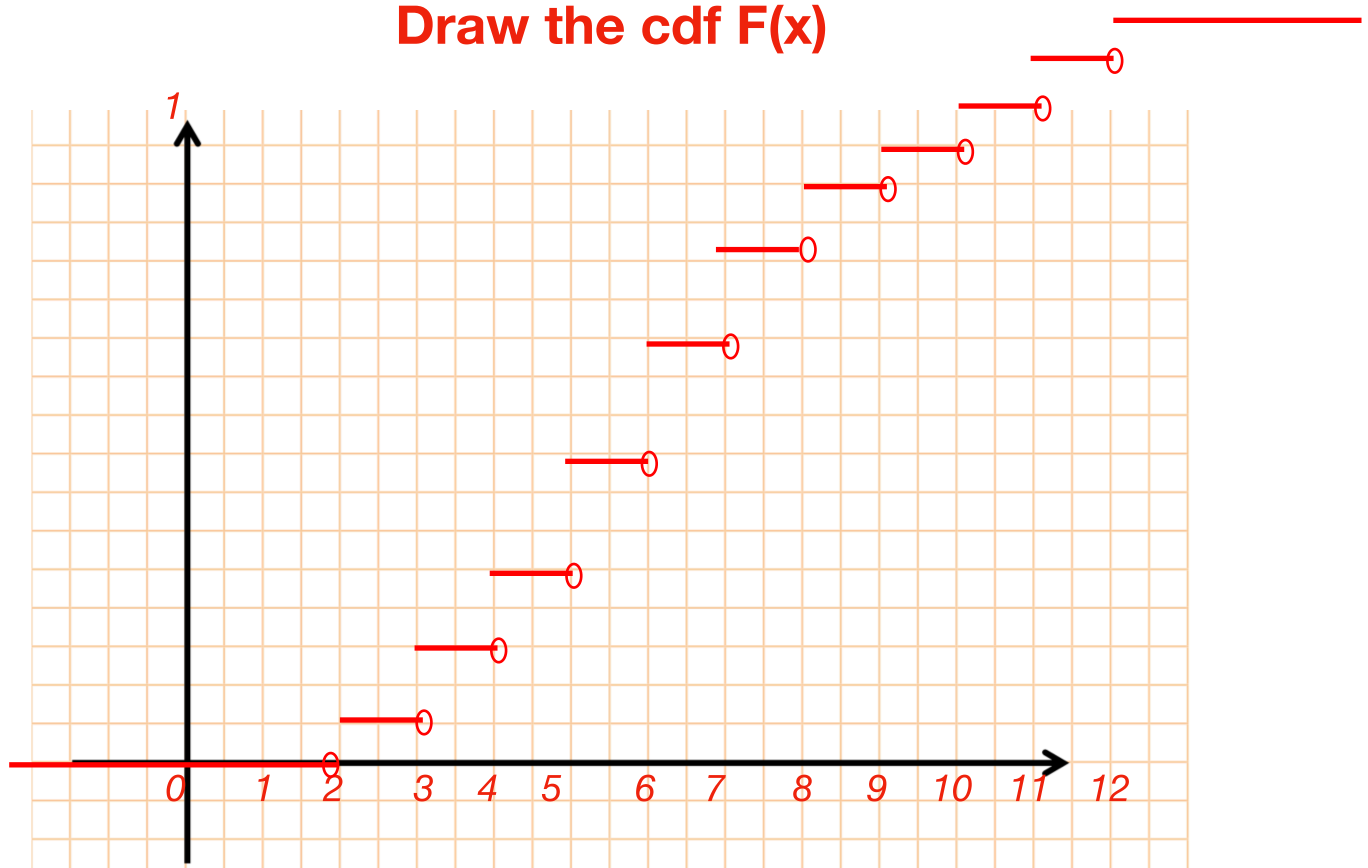
Example *Let X be the sum of the numbers obtained by rolling 2 fair dice.*

p.m.f.

x	2	3	4	5	6	7	8	9	10	11	12
P(X=x)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$$F(x) = \begin{cases} 0, & \text{if } x < 2 \\ 1/36, & \text{if } 2 \le x < 3 \\ 3/36, & \text{if } 3 \le x < 4 \\ 6/36, & \text{if } 4 \le x < 5 \\ 10/36, & \text{if } 5 \le x < 6 \\ 15/36, & \text{if } 6 \le x < 7 \\ 21/36, & \text{if } 7 \le x < 8 \\ 26/36, & \text{if } 8 \le x < 9 \\ 30/36, & \text{if } 9 \le x < 10 \\ 33/36, & \text{if } 10 \le x < 11 \\ 35/36, & \text{if } 11 \le x < 12 \\ 1, & \text{if } 12 \le x \end{cases}$$

Draw the cdf F(x)



4.2 Discrete Random Variables

Given p.m.f., how to obtain c.d.f.?

Example The distribution of the number X of mortgages approved per week at the local branch office of a bank is given below:

Approved per week	0	1	2	3	4	5	6
Probability	0.1	0.1	0.2	0.3	0.15	0.1	0.05

1). What is the probability that on a given week fewer than 4 home mortgages has been approved?

2). What is the probability that on a given week more than 2 but no more than 5 home mortgages had been approved

3). Draw the cdf of random variable X. .

4.2 Discrete Random Variables

Given c.d.f., how to obtain p.m.f.?

Example If the cumulative distribution function of a random variable X is

$$F(x) = \begin{cases} 0, & \text{if } x < 0; \\ 0.1, & \text{if } 0 \le x < 1; \\ 0.2, & \text{if } 1 \le x < 2; \\ 0.4, & \text{if } 2 \le x < 3; \\ 0.7, & \text{if } 3 \le x < 4; \\ 0.85, & \text{if } 4 \le x < 5; \\ 0.95, & \text{if } 5 \le x < 6; \\ 1, & \text{if } 6 \le x; \end{cases}$$

Find the probability mass function of X.

Important The difference of values of F between consecutive intervals gives the p.m.f.

$$F(x) = \begin{cases} 0, & \text{if } x < 0; \\ 0.1, & \text{if } 0 \le x < 1; \\ 0.2, & \text{if } 1 \le x < 2; \\ 0.4, & \text{if } 2 \le x < 3; \\ 0.7, & \text{if } 3 \le x < 4; \\ 0.85, & \text{if } 4 \le x < 5; \\ 0.95, & \text{if } 5 \le x < 6; \\ 1, & \text{if } 6 \le x; \end{cases}$$

$0.1 - 0 = 0.1$ for $x = 0$

$0.2 - 0.1 = 0.1$ for $x = 1$

$0.4 - 0.2 = 0.2$ for $x = 2$

$0.7 - 0.4 = 0.3$ for $x = 3$

$0.85 - 0.7 = 0.15$ for $x = 4$

$0.95 - 0.85 = 0.1$ for $x = 5$

$1 - 0.95 = 0.05$ for $x = 6$

4.3 Expected Value

Definition *If X is a discrete random variable having a probability mass function p_X, the **expectation** or the **expected value** of X, denoted by $E(X)$ or μ_X, is defined by*

$$E(X) = \sum_x x p_X(x).$$

Note that X is random variable , but E(X) is a deterministic number.

Commonly used notation:
Use U, V, X, Y, Z upper case of letters to denote random variables (for they are actually functions) and use $u, v, \ldots$ lower case of letters to denote values of random variables (values of random variables are just real numbers).

4.3 Expected Value

Interpretations of Expectation

See [Ross, pp. 125 – 126, 128]

(i) Weighted average of possible values that X can take on. Weights here are the probability that X assumes it.

(ii) Expectation is a measure of the central location of the random variable X.

(iii) Expectation is the average value of a random variable over a large number of experiments.

Example Suppose X takes only two values 0 and 1 with

$$P(X=0)=1-p \quad \text{and} \quad P(X=1)=p.$$

We call this random variable, a **Bernoulli** random variable of parameter p. And we denote it by $X \sim Be(p)$.

$$E(X)=0\times(1-p)+1\times p=p.$$

4.3 Expected Value

Example Let X denote the number obtained when a fair die is rolled. Then, $E(X) = 3.5$.

x	1	2	3	4	5	6
$p(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Solution

4.3 Expected Value

Example A newly married couple decides to continue to have children until they have one of each sex. If the events of having a boy and a girl are independent and equiprobable, how many children should this couple expect?

Solution

4.4 Expectation of a Function of a Random Variable

Class Discussion Let X be a random variable that takes values $-1, 0$ and 1 with probabilities:

$$P(X=-1)=0.2, \quad P(X=0)=0.5 \quad P(X=1)=0.3.$$

Define a new random variable $Y = X^2$.

What is the probability of Y=1?

Choice A	Choice B	Choice C
0.2	0.3	0.5

4.4 Expectation of a Function of a Random Variable

Given X, we are often interested about $g(X)$ and $E[g(X)]$. How do we compute $E[g(X)]$? One way is to find the probability mass function of $g(X)$ first and proceed to compute $E[g(X)]$ by definition.

Example Let X be a random variable that takes values $-1, 0$ and 1 with probabilities:

$$P(X = -1) = 0.2, \quad P(X = 0) = 0.5 \quad P(X = 1) = 0.3.$$

We are interested to compute $E(X^2)$.

Solution

4.4 Expectation of a Function of a Random Variable

Theorem *If X is a discrete random variable that takes values x_i, $i \geq 1$, with respective probabilities $p_X(x_i)$, then for any real value function g*

$$\begin{aligned} E[g(X)] &= \sum_i g(x_i) p_X(x_i) \quad \textit{or equivalently} \\ &= \sum_x g(x) p_X(x) \end{aligned}$$

Proof Group together all the terms in $\sum_i g(x_i)p(x_i)$ having the same value of $g(x_i)$. Suppose y_j, $j \geq 1$, represent the different values of $g(x_i)$, $i \geq 1$.

Then, grouping all the $g(x_i)$ having the same value gives

$$\sum_i g(x_i)p(x_i) = \sum_j \sum_{i:g(x_i)=y_j} g(x_i)p(x_i) \quad = \sum_j \sum_{i:g(x_i)=y_j} y_j p(x_i)$$

$$= \sum_j y_j \sum_{i:g(x_i)=y_j} p(x_i) \quad = \sum_i y_j P(g(X) = y_j) \quad = E[g(X)].$$

4.4 Expectation of a Function of a Random Variable

Example Let X be a random variable that takes values $-1, 0$ and 1 with probabilities:

$$P(X=-1)=0.2, \quad P(X=0)=0.5 \quad P(X=1)=0.3.$$

We are interested to compute $E(X^2)$.

Solution

Definition Take $g(x)=x^2$. Then,

$$E(X^2)=\sum_x x^2 p_X(x),$$

is called the **second moment** of X.

Moments are not random.

In general, for $k \geq 1$, $E\left(X^k\right)$ is called the k**-th moment** of X.

4.4 Expectation of a Function of a Random Variable

Definition Let $\mu = E(X)$, and take $g(x) = (x - \mu)^k$, then

$$E(X-\mu)^k \xrightarrow{\text{this notation means}} E[(X-\mu)^k]$$

is called the kth **central** moment.

Not $[E(X-\mu)]^k$

Remark (a) The expected value of a random variable X, $E(X)$ is also referred to as the **first moment** or the **mean** of X.

(b) The first central moment is 0.

(c) The second central moment, namely,

$$E(X-\mu)^2 \xrightarrow{\text{this notation means}} E[(X-\mu)^2]$$

is called the **variance** of X.

Not $[E(X-\mu)]^2$

4.4 Expectation of a Function of a Random Variable

Theorem *Let a and b be constants, then*

$$E[aX+b] = aE(X)+b.$$

Proof Apply $g(x) = ax+b$ in previous theorem, we have

$$\begin{aligned} E[aX+b] &= \sum_x [ax+b]p(x) \\ &= \sum_x [axp(x)+bp(x)] \\ &= a\sum_x xp(x) + b\sum_x p(x) \\ &= aE(X)+b. \end{aligned}$$

4.4 Expectation of a Function of a Random Variable

Theorem (Tail Sum Formula for Expectation).
For nonnegative integer-valued random variable X (that is, X takes values $0, 1, 2, \ldots$),

$$E(X) = \sum_{k=1}^{\infty} P(X \geq k) = \sum_{k=0}^{\infty} P(X > k).$$

Proof Consider the following triangularization:

$$\begin{aligned}
\sum_{k=1}^{\infty} P(X \geq k) &= P(X=1) + P(X=2) + P(X=3) + P(X=4) + \cdots && P(X \geq 1) \\
& + P(X=2) + P(X=3) + P(X=4) + \cdots && P(X \geq 2) \\
& + P(X=3) + P(X=4) + \cdots && P(X \geq 3) \\
& + P(X=4) + \cdots && \vdots \\
& + \cdots && \vdots \\
&= P(X=1) + 2P(X=2) + 3P(X=3) + 4P(X=4) + \cdots && \vdots \\
&= E(X).
\end{aligned}$$

4.5 Variance and Standard Deviation

Definition *If X is a random variable with mean μ, then the* ***variance*** *of X, denoted by* var(X)*, is defined by*

$$\operatorname{var}(X) = E(X-\mu)^2.$$

It is a measure of scattering (or spread) of the values of X.

Definition *The* ***standard deviation*** *of X, denoted by σ_X or* SD(X)*, is defined as*

$$\sigma_X = \sqrt{\operatorname{var}(X)}.$$

μ_X is expectation

σ_X is standard deviation

Interpretaion Variance measures the spread of a random variable X.

If values of X near its mean μ_X are very likely and values further away from μ_X have very small probability, then the distribution of X will be closely concentrated around μ_X. In this case, the spread of the distribution of X is small. On the other hand, if values of X some distance from its mean μ_X are likely, the spread of the distribution of X will be large.

4.5 Variance and Standard Deviation

Theorem $\operatorname{var}(X) = E(X^2) - [E(X)]^2.$

Proof

4.5 Variance and Standard Deviation

Remark (1) Note that $\mathrm{var}(X) \geq 0$. (Why?)

(2) $\mathrm{var}(X) = 0$ if and only if X is a **degenerate** random variable (that is, the random variable taking only one value, its mean).

(3) It follows from the formula that

$$E(X^2) \geq [E(X)]^2 \geq 0.$$

Example Calculate $\mathrm{var}(X)$ if X represents the outcome when a fair die is rolled.

Solution

4.5 Variance and Standard Deviation

Theorem

(i) $\mathrm{var}(aX+b) = a^2\,\mathrm{var}(X)$.

(ii) $\mathrm{SD}(aX+b) = |a|\,\mathrm{SD}(X)$.

here a and b are constants.

Proof

4.5 Variance and Standard Deviation

Class Discussion

Which of the following ones is the 2nd central moment of the random variable X?

Choice A

$E(X^2)$

Choice B

$\mathrm{Var}(X)$

Choice C

$(E(X))^2$

4.5 Variance and Standard Deviation

Example

Consider the probability distribution for the returns on stock A and B provided below:

	X	Y
Probability	Return on Stock A	Return on Stock B
0.2	1%	10%
0.3	2%	6%
0.3	3%	2%
0.2	4%	-2%

Find the expected return on stock A and stock on B:

Solution

4.5 Variance and Standard Deviation

Example (continue)

Find the variance and standard deviation of return on stock A and stock on B:

Solution

4.6 Discrete Random Variables arising from Repeated Trials

We study a mathematical model for repeated trials:

(1) Each trial results in whether a particular event occurs or doesn't. Occurrence of this event is called **success**, and non-occurrence called **failure**. Write $p := P(\text{success})$, and $q := 1 - p = P(\text{failure})$.

nature of trial	meaning of success	meaning of failure	probabilities p and q
Flip a fair coin	head	tail	0.5 and 0.5
Roll a fair die	six	non-six	$1/6$ and $5/6$
Roll a pair of fair dice	double six	not double six	$1/36$ and $35/36$
Birth of a child	girl	boy	0.487 and 0.513
Pick an outcome	in A	not in A	$P(A)$ and $1 - P(A)$

(2) Each trial with success probability p, failure with $q = 1 - p$;

(3) We repeat the trials independently.

Such trials are called **Bernoulli**(p) **trials**.

4.6 Discrete Random Variables arising from Repeated Trials

Bernoulli random variable

We only perform the experiment once, and define

$$X = \begin{cases} 1 & \text{if it is a success} \\ 0 & \text{if it is a failure} \end{cases}.$$

Here

$$P(X=1) = p, \quad P(X=0) = 1-p$$

and

$$E(X) = p, \quad \text{var}(X) = p(1-p).$$

Denoted by Be(p)

A random variable has a distribution.
Bernoulli random variable's distribution is called Bernoulli distribution.

4.6 Discrete Random Variables arising from Repeated Trials

Binomial random variable

We perform the experiment (under identical conditions and independently) n times and define

$$X = \text{ number of successes in } n \text{ Bernoulli}(p) \text{ trials.}$$

Therefore, X takes values $0, 1, 2, \ldots, n$. In fact, for $0 \leq k \leq n$,

$$P(X = k) = \binom{n}{k} p^k q^{n-k}.$$

the p.m.f. is derived similarly as the example on slide 59 of Chapter 3

Denoted by Bin(n, p)

Binomial random variable

Binomial distribution

4.6 Discrete Random Variables arising from Repeated Trials

Theorem *If X ~ Bin(n, p), then* $E(X) = np, \quad \mathrm{var}(X) = np(1-p).$

Proof

4.6 Discrete Random Variables arising from Repeated Trials

Geometric random variable

Define the random variable X = number of Bernoulli(p) trials required to obtain the first success.

(Note, the trial leading to the first success is included.) Here, X takes values $1, 2, 3, \ldots$ and so on. In fact, for $k \geq 1$,

$$P(X = k) = pq^{k-1}.$$

And

$$E(X) = \frac{1}{p}, \quad \mathrm{var}(X) = \frac{1-p}{p^2}.$$

Denoted by Geom(p)

Geometric random variable

Geometric distribution

4.6 Discrete Random Variables arising from Repeated Trials

Geometric random variable

Another version of Geometric distribution:

$$X' = \text{number of failures in the Bernoulli}(p) \text{ trials in order to obtain the first success.}$$

Here

$$X = X' + 1.$$

Hence, X' takes values $0, 1, 2, \ldots$ and

$$P(X' = k) = pq^k, \qquad \text{for } k = 0, 1, \ldots.$$

And

$$E(X') = \frac{1-p}{p}, \quad \text{var}(X') = \frac{1-p}{p^2}.$$

Conventionally, the Geometric distribution refers to X rather than X'

4.6 Discrete Random Variables arising from Repeated Trials

Negative Binomial random variable

Define the random variable

$$X = \text{number of Bernoulli}(p) \text{ trials required to obtain } r \text{ success.}$$

Here, X takes values $r, r+1, \ldots$ and so on. In fact, for $k \geq r$,

$$P(X = k) = \binom{k-1}{r-1} p^r q^{k-r}.$$

And $$E(X) = \frac{r}{p}, \quad \text{var}(X) = \frac{r(1-p)}{p^2}.$$

Denoted by NB(r, p)

Take note that $\text{Geom}(p) = \text{NB}(1, p)$.

Why called negative binomial ? $\binom{k-1}{r-1} = (-1)^{r-1}\binom{-(k-r+1)}{r-1}$

4.6 Discrete Random Variables arising from Repeated Trials

Example A gambler makes a sequence of 1-dollar bets, betting each time on black at roulette at Las Vegas. Here a success is winning 1 dollar and a failure is losing 1 dollar. Since in American roulette the gambler wins if the ball stops on one of 18 out of 38 positions and loses otherwise, the probability of winning is $p = 18/38 = 0.474$.

Example In a small town, out of 12 accidents that occurred in 1986, *at least four* happened on Friday the 13th. Is this a good reason for a superstitious person to argue that Friday the 13th is inauspicious?

Suppose the probability that each accident occurs on Friday the 13th is $1/30$, just as on any other day. Then the probability of at least four accidents on Friday the 13th is

$$1 - \sum_{i=0}^{3} \binom{12}{i} \left(\frac{1}{30}\right)^{i} \left(\frac{29}{30}\right)^{12-i} \approx 0.000493.$$

Since this probability is small, this is a good reason for a superstitious person to argue that Friday the 13th is inauspicious.

4.6 Discrete Random Variables arising from Repeated Trials

Example The geometric distribution plays an important role in the theory of queues, or waiting lines. For example, suppose a line of customers waits for service at a counter. It is often assumed that, in each small time unit, either 0 or 1 new customers arrive at the counter. The probability that a customer arrives is p and that no customer arrives is $q = 1 - p$. Then the time T until the next arrival has a geometric distribution. It is natural to ask for the probability that no customer arrives in the next k time units, that is, for $P(T > k)$.

Solution

This is given by
$$P(T > k) = \sum_{j=k+1}^{\infty} q^{j-1} p = q^k(p + qp + q^2 p + \cdots) = q^k.$$

This probability can also be found by noting that we are asking for no successes (i.e., arrivals) in a sequence of k consecutive time units, where the probability of a success in any one time unit is p. Thus, the probability is just q^k.

4.6 Discrete Random Variables arising from Repeated Trials

Example

Ten students are asked to randomly pick one number between 0 and 9 inclusively. Let X be the random variable of the number of students who pick the number "8".

Find the probability that more than 1 student pick the number "8"

Solution

4.6 Discrete Random Variables arising from Repeated Trials

Example A communication system consists of n components, each of which will, independently, function with probability p. The total system will be able to operate effectively if at least one-half of its components function.

For what values of p is a 5-component system more likely to operate effectively than a 3-component system?

Solution

4.7 Poisson Random Variable

A random variable X is said to have a **Poisson** distribution with parameter λ if X takes values $0, 1, 2, \ldots$ with probabilities given as:

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}, \qquad \text{for } k = 0, 1, 2, \ldots.$$

This defines a probability mass function, since

$$\sum_{k=0}^{\infty} P(X = k) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1.$$

Notation: $X \sim \text{Poisson}(\lambda)$.

4.7 Poisson Random Variable

Theorem If $X \sim \text{Poisson}(\lambda)$, $\quad E(X) = \lambda, \quad \text{var}(X) = \lambda.$

Proof

4.7 Poisson Random Variable

The Poisson random variable has a tremendous range of application in diverse areas because It can be used to determine the probability of counts of the occurrence of an event over time (or space).

Example (of Poisson distribution)

1. The number of traffic accidents occurring on a highway in a day.
2. Crashes of a computer network per week.
3. The number of people joining a line in an hour.
4. The number of customers arrived per day.
5. The number of goals scored in a hockey game.
6. The number of typos per page of an essay.

Sounds similar to a Binomial distribution ?

4.7 Poisson Random Variable

Poission random variable can be used as an approximation for a binomial random variable with parameters (n, p) when n is large and p is small enough so that np is of moderate size. To see this, suppose X is a binomial random variable with parameters (n, p) and let $\lambda = np$. Then

$$\begin{aligned} P(X = k) &= \frac{n!}{(n-k)!k!} p^k (1-p)^{n-k} \\ &= \frac{n!}{(n-k)!k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{n^k} \frac{\lambda^k}{k!} \frac{(1-\lambda/n)^n}{(1-\lambda/n)^k} \end{aligned}$$

Note that for n large and λ moderate,

$$\left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda}, \frac{n(n-1)\cdots(n-k+1)}{n^k} \approx 1 \text{ and } \left(1 - \frac{\lambda}{n}\right)^k \approx 1.$$

Hence for n large and λ moderate,

$$P(X = k) \approx e^{-\lambda} \frac{\lambda^k}{k!}.$$

Usually if $n > 20$ and $np < 15$ then approximation is valid

$\text{Bin}(n, p) \approx \text{Poisson}(np)$

4.7 Poisson Random Variable

Remark In other words, if n independent trials, each of which results in a success with probability p are performed, then when n is large and p is small enough to make np moderate, the number of successes occurring is approximately aPoisson random variable with parameter $\lambda = np$.

(i) number of misprints on a page;

(ii) number of people in a community living to 100 years;

(iii) number of wrong telephone numbers that are dialed in a day;

(iv) number of people entering a store on a given day;

Each of the preceding, and numerous other random variables, are approximately Poisson for the same reason – because of the Poisson approximation to the binomial. ***when n becomes large.***

4.7 Poisson Random Variable

Class Discussion

Which of the following random variables has an infinite range?

Choice A

$\mathrm{Be}(p)$

Choice B

$\mathrm{Poisson}(\lambda)$

Choice C

$\mathrm{Bin}(n, p)$

4.7 Poisson Random Variable

Example Suppose that the number of typographical errors on a page of a book has a Poisson distribution with parameter $\lambda = \frac{1}{2}$. Calculate the probability that there is at least one error on a page.

Solution

Example Suppose that the probability that an item produced by a certain machine will be defective is 0.1. Find the probability that a sample of 10 items will contain at most 1 defective item.

Solution

4.7 Poisson Random Variable

Example

Suppose during a particular minute of the day the n = 2000000 people serviced in a particular telephone service area decide independently of each other whether to place an emergency call to 911 or not. Each person has probability p=.000005 of doing so. Let T be the actual random number of 911 callers in that minute. Find the probability P(T=10).

Solution

4.7 Poisson Random Variable

Example During a laboratory experiments, the average number of radioactive parties passing through a counter in 1 millisecond is 4. Suppose that the number of particles passing through the counter follows a Poisson distribution. What is the probability that 6 particles enter the counter in a given millisecond?

Solution

4.7 Poisson Random Variable

Example

From June 1944 to March 1945 during World War II, Germany launched a total of 9,251 V-1 flying bombs — "buzz" bombs — against England. Of these, only 2,419 made it to their intended target areas and, of these, 537 struck South London.

Shortly after World War II, a British statistician named *R.D. Clarke* took a 12 km x 12 km heavily bombed region of South London, and sliced it up in to a grid. In all, he divided it into 576 squares (or regions), each about the size of 25 city blocks. Next, he counted the number of regions with 0 bombs dropped, 1 bomb dropped, 2 bombs dropped, and so on.

The red dots show where the flying bombs landed in South London

Clarke then showed that the pattern of "hits" would follow a Poisson distribution!

Now based on his result, find the probability that a randomly selected region was hit exactly twice.

First, since 537 bombs struck 576 regions, the mean number of hits per region is

$$\lambda = \frac{537}{576} = 0.9323$$

Let X be the rv of the number of hits in the selected region. Thus,

$$X \sim Poisson(0.9323)$$

and the required probability of the selected region being hit exactly twice is

$$P(X = 2) = \frac{(0.9323)^2 e^{-0.9323}}{2!} = 0.1711.$$

4.7 Poisson Random Variable

Example

We are studying the earthquakes in California with a reading over 6.7 on the Richter scale. Suppose that ***on average***, there are 1.5 earthquakes with a reading over 6.7 on the Richter scale in California per year.

In this case, $\lambda = 1.5$ is called the rate of the occurrences of the earthquakes above 6.7 on the Richter scale, and the time unit is 1 year.

Let X be the rv of the number of earthquakes above 6.7 on the Richter scale in the upcoming year, then

$$X \sim Poisson(1.5).$$

Thus, the probability that there will be 5 earthquakes with a reading over 6.7 on the Richter scale in the upcoming year is

$$P(X = 5) = \frac{(1.5)^5 e^{-1.5}}{5!} = 0.0141.$$

4.7 Poisson Random Variable

In the previous example

How about the probability that there will be 5 earthquakes with a reading over 6.7 on the Richter scale ***in the next 4 years***?

To answer this question, we need the following result:

When we study the count of occurrences of an event over a period of t units of time with the rate λ of the occurrences per unit time, use

$$Y_t \sim Poisson(\lambda t),$$

where Y_t is the random variable of the count of occurrences of the event over a period of time t.

Thus, the required probability is

$$P(Y_4 = 5) = \frac{(\lambda t)^5 e^{-(\lambda t)}}{5!} = \frac{(1.5 \times \mathbf{4})^5 e^{-(1.5 \times \mathbf{4})}}{5!} = 0.1606.$$

4.7 Poisson Random Variable

Theorem

If the number of occurrence of an event in unit time (or space) followings Poisson distribution with rate λ, then the number of occurrence in t units time (or spaces) followings Poisson(λt)

Proof is not trivial……

4.7 Poisson Random Variable

Example

The average number of homes sold by the Centaline agency is 2 homes per day (Poisson distribution). What is the probability that exactly 10 homes will be sold by Centaline in the next 30 days?

Solution

4.8 Hypergeometric Random Variable

Suppose that we have a set of N balls, of which m are red and $N-m$ are blue. We choose n of these balls, *without replacement*, and define X to be the number of red balls in our sample. Then

$$P(X=x)=\frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}},$$

for $x=0,1,\ldots,\ \min(m,n)$

A random variable whose probability mass function is given as the above equation for some values of n, N, m is said to be a **hypergeometric** random variable, and is denoted by $H(n,N,m)$. Here

$$E(X)=\frac{nm}{N},\quad \mathrm{var}(X)=\frac{nm}{N}\left[\frac{(n-1)(m-1)}{N-1}+1-\frac{nm}{N}\right].$$

4.8 Hypergeometric Random Variable

Example A purchaser of electrical components buys them in lots of size 10. It is his policy to inspect 3 components randomly from a lot and to accept the lot only if all 3 are nondefective. If 30 percent of the lots have 4 defective components and 70 percent have only 1, what proportion of lots does the purchaser reject?

Solution

4.9 Expected Value of Sum of Random Variables

For a random variable X, let $X(s)$ denote the value of X when $s \in S$ is the outcome. Now, if X and Y are both random variables, then so is their sum. That is, $Z = X + Y$ is also a random variable. Moreover, $Z(s) = X(s) + Y(s)$.

Recall that a R.V. is a map from sample space to the real line.

Example Suppose that an experiment consists of flipping a coin 5 times, with the outcome being the resulting sequence of heads and tails. Let X be the number of heads in the first 3 flips and Y the number of heads in the final 2 flips. Let $Z = X + Y$. Then for the outcome $s = (h, t, h, t, h)$,

$$X(s) = 2, \quad Y(s) = 1, \quad Z(s) = X(s) + Y(s) = 3.$$

For the outcome $s = (h, h, h, t, h)$,

$$X(s) = 3, \quad Y(s) = 1, \quad Z(s) = X(s) + Y(s) = 4.$$

4.9 Expected Value of Sum of Random Variables

Let $p(s) = P(\{s\})$ be the probability that s is the outcome of the experiment.

Theorem

$$E[X] = \sum_{s \in S} X(s)p(s).$$

Proof. Suppose that the distinct values of X are $x_j, i \geq 1$. For each i, let S_i be the event that X is equal to x_i. That is, $S_i = \{s : X(s) = x_i\}$. Then,

$$\begin{aligned}
E[X] &= \sum_i x_i P\{X = x_i\} \\
&= \sum_i x_i P(S_i) \\
&= \sum_i x_i \sum_{s \in S_i} p(s) \\
&= \sum_i \sum_{s \in S_i} x_i p(s) \\
&= \sum_i \sum_{s \in S_i} X(s)p(s) \\
&= \sum_{s \in S} X(s)p(s).
\end{aligned}$$

4.9 Expected Value of Sum of Random Variables

Example Suppose that two independent flips of a coin that comes up heads with probability p are made, and let X denote the number of heads obtained. Now

$$\begin{aligned} P(X=0) &= P(t,t) = (1-p)^2, \\ P(X=1) &= P(h,t)+P(t,h) = 2p(1-p), \\ P(X=2) &= P(h,h) = p^2 \end{aligned}$$

It follows from the definition that

$$E[X] = 0\times(1-p)^2 + 1\times 2p(1-p) + 2\times p^2 = 2p$$

which agrees with

$$\begin{aligned} E[X] &= X(h,h)p^2 + X(h,t)p(1-p) + X(t,h)(1-p)p + X(t,t)(1-p)^2 \\ &= 2p^2 + p(1-p) + (1-p)p \\ &= 2p. \end{aligned}$$

4.9 Expected Value of Sum of Random Variables

Theorem *For random variables $X_1, X_2, \ldots, X_n$,* $$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E\left[X_i\right].$$

Proof

4.9 Expected Value of Sum of Random Variables

Example Find the expected total number of successes that result from n trials when trial i is a success with probability $p_i, i = 1, \ldots, n$.

Solution

4.9 Expected Value of Sum of Random Variables

Example *Suppose there are n identical and independent trials, each trial yields success with the same probability p.*

Let

$$X_i = \begin{cases} 1, & \text{if trial } i \text{ is a success} \\ 0, & \text{if trial } i \text{ is a failure} \end{cases}.$$

Show that $X_1 + X_2 + \cdots + X_n \sim \mathrm{Bin}(n, p)$

Solution

4.9 Expected Value of Sum of Random Variables

Example For $X \sim \text{Bin}(n, p)$, find $\text{var}(X)$.

Solution