Chapter 3. Discrete random variables and probability distributions.

Defn: A random variable (r.v.) is a function that takes a sample point in S and maps it to it a real number.
That is, a r.v. Y maps the sample space (its domain) to the real number line (its range), $Y : S \to \mathbb{R}$.

Example: In an experiment, we roll two dice. Let the r.v. Y = the sum of two dice.
The r.v. Y as a function:
$(1, 1) \to 2$
$(1, 2) \to 3$
$\vdots$
$(6, 6) \to 12$

Discrete random variables and their distributions

Defn: A discrete random variable is a r.v. that takes on a finite or countably infinite number of different values.

Defn: The probability mass function (pmf) of a discrete r.v Y is a function that maps each possible value of Y to its probability. (It is also called the probability distribution of a discrete r.v. Y).

$p_Y : Range(Y) \to [0, 1]$.

$p_Y(v) = \mathbb{P}(Y = v)$, where 'Y=v' denotes the event $\{\omega : Y(\omega) = v\}$.

Notation: We use capital letters (like Y) to denote a random variable, and a lowercase letters (like v) to denote a particular value the r.v. may take.

Cumulative distribution function

Defn: The cumulative distribution function (cdf) is defined as $F(v) = \mathbb{P}(Y \leq v)$.

$p_Y : Range(Y) \to [0, 1]$.

Example: We toss two fair coins. Let Y be the number of heads. What is the pmf of Y? What is the cdf?

Quiz

X a random variable.

values of X:	1	3	5	7
cdf F(a):	.5	.75	.9	1

What is $P(X \leq 3)$?

(A) .15
(B) .25
(C) .5
(D) .75

Quiz

X a random variable.

values of X:	1	3	5	7
cdf F(a):	.5	.75	.9	1

What is $P(X = 3)$?

(A) .15
(B) .25
(C) .5
(D) .75

Quiz

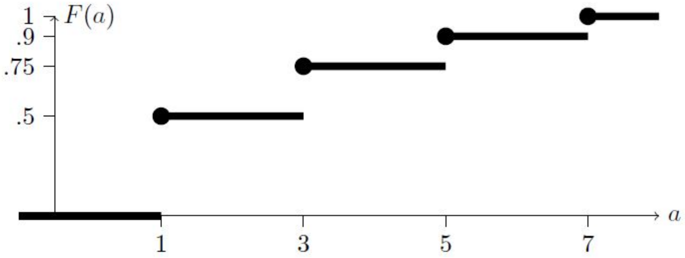
X a random variable.

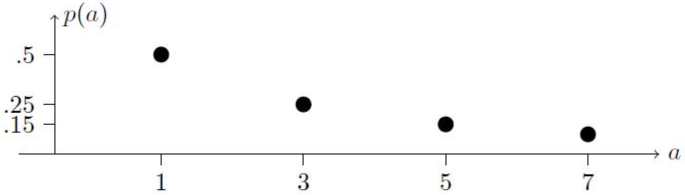
values of X:	1	3	5	7
cdf F(a):	.5	.75	.9	1

What is $P(X < 3)$?

(A) .15
(B) .25
(C) .5
(D) .75

CDF and PMF





Example

In an experiment, we roll two fair dice. Let the r.v. Y = the sum of dice. What is the pmf of Y?

Properties of probability distributions

Theorem For any discrete probability distribution, the following is true:

1. $0 \leq p(y) \leq 1$.
2. $\sum_y p(y) = 1$, where the summation is over all values y with non-zero probability.

Expected value of a r.v.

Defn: Let Y be a discrete r.v. with probability distribution $p(y)$. Then the expected value of Y, denoted $\mathbb{E}(Y)$, is:

$$\mathbb{E}(Y) = \sum_y y p(y).$$

Note: $\mathbb{E}(Y)$ exists if this sum is absolutely convergent (if $\sum_y |y| p(y) < \infty$).

$\mathbb{E}(Y)$ is often called the mean or average of Y and denoted μ.

Example: Roll a die many times. What is the average value?

Quiz

We roll two dice. You win \$1000 if the sum is 2 and lose \$100 otherwise. How much do you expect to win on average per trial?

(A) -\$13.11
(B) \$69.44
(C) -\$69.44
(D) \$13.11
(E) None of the above

Expected value of a function of a r.v.

Theorem: If Y is a discrete r.v. with probability function $p(y)$ and $g(Y)$ is a real-valued function of Y, then

$$\mathbb{E}[g(Y)] = \sum_y g(y)p(y).$$

Proof: Suppose $g(Y)$ take values $g_1, \ldots, g_m$. Then by definition

$$\begin{aligned}
\mathbb{E}[g(Y)] &= \sum_{i=1}^{m} g_i \mathbb{P}(g(Y) = g_i) \\
&= \sum_{i=1}^{m} g_i \sum_{y:g(y)=g_i} p(y) \\
&= \sum_{i=1}^{m} \sum_{y:g(y)=g_i} g_i p(y) = \sum_y g(y)p(y)
\end{aligned}$$

$\square$

Example: Toss two fair coins. If Y is the number of heads, what is $\mathbb{E}(Y^2)$?

Properties of expected value

Theorem Let a and b be constant and Y is a r.v. Then $\mathbb{E}[aY + b] = a\mathbb{E}[Y] + b$.

Theorem Let $Y_1, Y_2, \ldots Y_k$ be r.v.'s, then

$$\mathbb{E}[Y_1 + Y_2 + \ldots + Y_k] = \mathbb{E}Y_1 + \mathbb{E}Y_2 + \ldots + \mathbb{E}Y_k.$$

Proof:

Corollary

$$\mathbb{E}[g_1(Y) + g_2(Y) + \ldots + g_k(Y)] = \mathbb{E}g_1(Y) + \mathbb{E}g_2(Y) + \ldots + \mathbb{E}g_k(Y).$$

Example

Suppose that n people are sitting around a table, and that everyone at the table got up, ran around the room, and sat back down randomly (i.e., all seating arrangements are equally likely).

What is the expected value of the number of people sitting in their original seat?

Petersburg Paradox

Consider a series of flips of a fair coin. A player will receive \$2 if the first head occurs on flip 1, \$4 if it occurs on flip 2, \$8 if on flip 3, and so on.
(1) What is the player's expected earning?
(2) How much one should pay for the right to play this game?

H
TH
TTH
TTTH
TTTTH
...

Variance and standard deviation

Defn: The variance is the expected squared difference between a random variable and its mean,

$$\mathrm{Var}(Y) = \mathbb{E}[(Y-\mu)^2],$$

where $\mu = \mathbb{E}(Y)$.
Defn: The standard deviation of Y is $\sigma = \sqrt{\mathrm{Var}(Y)}$

Formula for variance:

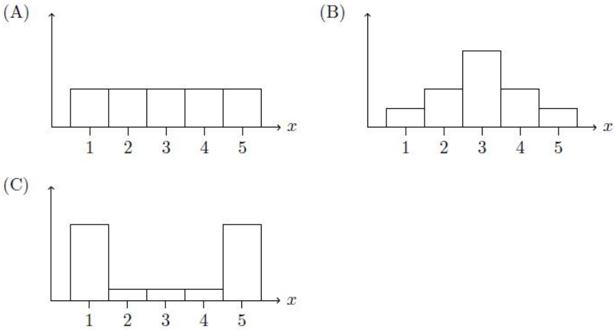
$$\mathrm{Var}(Y) = \mathbb{E}(Y^2) - \mu^2$$

Proof:

$$\begin{aligned} \mathrm{Var}(Y) &= \mathbb{E}(Y^2 - 2\mu Y + \mu^2) \\ &= \mathbb{E}(Y^2) - 2\mu\mathbb{E}(Y) + \mu^2 \\ &= \mathbb{E}(Y^2) - \mu^2. \end{aligned}$$

$\square$

Quiz



These graphs show the pmf for 3 random variables. Order them by size of standard deviation from biggest to smallest.

(A) ABC
(B) ACB
(C) BAC
(D) BCA
(E) CAB

Example

Suppose Y has pmf:

y	1	2	3	4	5
$p(y)$	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{4}{10}$	$\frac{2}{10}$	$\frac{1}{10}$

Find the expectation, variance and the standard deviation of Y.

Properties of the variance and standard deviation

If a and b are constants, then

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

$$\sigma(aX + b) = |a|\sigma(X).$$

Quiz

True or false: If $\mathrm{Var}(X) = 0$ then X is constant.

(A) True
(B) False

What about the variance of the sum? Independent r.v.

Defn: Two discrete r.v.'s X and Y are called independent if the events $\{X = x\}$ and $\{Y = y\}$ are independent for all possible values of x and y.

This is equivalent to the following
Theorem Two discrete r.v.'s X and Y are independent if and only if $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$ for all possible functions f and g.

Proof:

Theorem If r.v.'s X_1, X_2, ..., X_n are independent, then

$$\mathrm{Var}[X_1 + \ldots + X_n] = \mathrm{Var}X_1 + \ldots + \mathrm{Var}X_n$$

Proof:

Discrete r.v.

- Binomial
- Bernoulli
- Geometric
- Poisson
- Negative binomial
- Hypergeometric

Discrete r.v.

- Binomial
- Bernoulli
- Geometric
- Poisson
- Negative binomial
- Hypergeometric

Discrete r.v.

- Binomial
- Bernoulli
- Geometric
- Poisson
- Negative binomial
- Hypergeometric

Discrete r.v.

- Binomial
- Bernoulli
- Geometric
- Poisson
- Negative binomial
- Hypergeometric

Discrete r.v.

- Binomial
- Bernoulli
- Geometric
- Poisson
- Negative binomial
- Hypergeometric

Discrete r.v.

- Binomial
- Bernoulli
- Geometric
- Poisson
- Negative binomial
- Hypergeometric

Binomial Experiment

Example: Suppose 2% of all items produced from an assembly line are defective. We randomly sample 50 items and count how many are defective (and how many are not).

Defn: A binomial experiment is an experiment with the following characteristics:

1. There are a fixed number, denoted n, of identical trials.
2. Each trial results in one of two outcomes (which we denote "Success" or "Failure").
3. The probability of Success is constant across trials and denoted p. Hence $\mathbb{P}[\text{Failure}] = 1 - p$.
4. The trials are independent.

Binomial r.v. and binomial distribution

Example: Suppose 40% of students at a college are male. We select 10 students at random and count how many are male. Is this a binomial experiment?

Defn: The total number of successes in a binomial experiment is called a binomial r.v.

Defn: The probability mass function of a binomial random variables is called the binomial distribution.

Formula for the binomial distribution

If we have n trials, a typical sample point looks like:
SFFFSSFSFFF...
Of the n slots, suppose we have k successes.

1. How many ways to select k slots where to put S letter?

2. For any sample point containing k S's and $(n-k)$ F's, what is its probability?

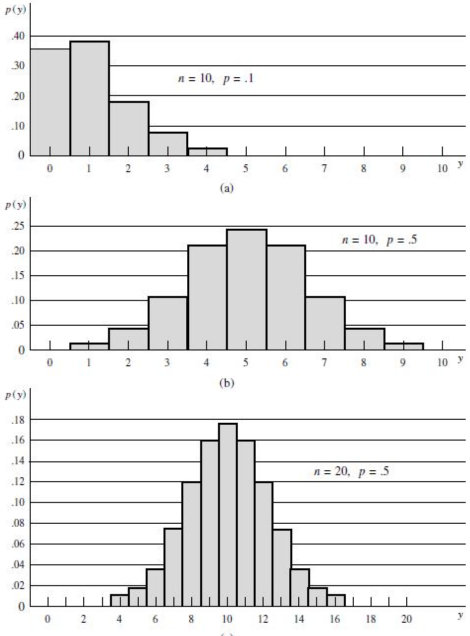
Corresponding to $Y = k$, we have $\binom{n}{k}$ sample points,
each with probability $p^k q^{n-k}$, where $q = 1 - p$.

So for any $k = 0, 1, \ldots, n$:

$$\mathbb{P}(Y = k) = \binom{n}{k} p^k q^{n-k}$$

This is the binomial probability distribution.

Pictures of binomial distribution



How to calculate binomial distribution

Example Suppose 2% of items produced from an assembly line are defective. If we sample 10 items, what is the probability that 2 or more are defective?

For many problems, we use tables or software to find binomial probabilities.

Table 1, Appendix 3 gives cumulative probabilities $\mathbb{P}(Y \leq a)$:

Quiz 9

40% of students in a college are male. 10 students are selected. What is the probability that 7 or fewer students are selected?

Table 1 Binomial Probabilities

Tabulated values are $P(Y \le a) = \sum_{y=0}^{a} p(y)$. (Computations are rounded at third decimal place.)

(a) $n = 5$

a	0.01	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95	0.99	a
0	.951	.774	.590	.328	.168	.078	.031	.010	.002	.000	.000	.000	.000	0
1	.999	.977	.919	.737	.528	.337	.188	.087	.031	.007	.000	.000	.000	1
2	1.000	.999	.991	.942	.837	.683	.500	.317	.163	.058	.009	.001	.000	2
3	1.000	1.000	1.000	.993	.969	.913	.812	.663	.472	.263	.081	.023	.001	3
4	1.000	1.000	1.000	1.000	.998	.990	.969	.922	.832	.672	.410	.226	.049	4

(b) $n = 10$

a	0.01	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95	0.99	a
0	.904	.599	.349	.107	.028	.006	.001	.000	.000	.000	.000	.000	.000	0
1	.996	.914	.736	.376	.149	.046	.011	.002	.000	.000	.000	.000	.000	1
2	1.000	.988	.930	.678	.383	.167	.055	.012	.002	.000	.000	.000	.000	2
3	1.000	.999	.987	.879	.650	.382	.172	.055	.011	.001	.000	.000	.000	3
4	1.000	1.000	.998	.967	.850	.633	.377	.166	.047	.006	.000	.000	.000	4
5	1.000	1.000	1.000	.994	.953	.834	.623	.367	.150	.033	.002	.000	.000	5
6	1.000	1.000	1.000	.999	.989	.945	.828	.618	.350	.121	.013	.001	.000	6
7	1.000	1.000	1.000	1.000	.998	.988	.945	.833	.617	.322	.070	.012	.000	7
8	1.000	1.000	1.000	1.000	1.000	.998	.989	.954	.851	.624	.264	.086	.004	8
9	1.000	1.000	1.000	1.000	1.000	1.000	.999	.994	.972	.893	.651	.401	.096	9

Using R to compute binomial probabilities

R is both a language and a software tool which is very popular among statisticians.

One can download and install R software from
https://cran.rstudio.com/,
and RStudio (tools for using R) from
https://www.rstudio.com/products/rstudio/download/

Then, we calculate $\mathbb{P}(Y \leq 7)$ by issuing command:
pbinom(7,10,0.4)

Answer: 0.9877054

Mean and variance of the binomial r.v.

Theorem: Let Y be a binomial r.v. with n trials and success probability p. Then,

1. $\mathbb{E}(Y) = np$
2. $\mathrm{Var}(Y) = np(1-p)$

Proof: (1)

$$\begin{aligned}
\mathbb{E}(Y) &= \sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k q^{n-k} \\
&= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k q^{n-k} \\
&= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} q^{n-k} \\
&= np \sum_{l=0}^{n-1} \frac{(n-1)!}{l!(n-l-1)!} p^l q^{n-l-1} \\
&= np.
\end{aligned}$$

Variance of the binomial r.v.

Proof: (2) $\mathrm{Var}(Y) = \mathbb{E}[Y^2] - (\mathbb{E}(Y))^2$.

But it is difficult to find $\mathbb{E}(Y^2)$ directly. So we use a trick.

1.
$$\begin{aligned}
\mathbb{E}[Y(Y-1)] &= \sum_{k=0}^{n} k(k-1)\frac{n!}{k!(n-k)!}p^k q^{n-k} \\
&= \cdots \\
&= n(n-1)p^2.
\end{aligned}$$

2.
$$\begin{aligned}
\mathbb{E}(Y^2) &= \mathbb{E}[Y(Y-1)] + \mathbb{E}(Y) \\
&= n(n-1)p^2 + np
\end{aligned}$$

3.
$$\begin{aligned}
\mathrm{Var}(Y) &= \mathbb{E}[Y^2] - (\mathbb{E}(Y))^2 \\
&= n(n-1)p^2 + np - n^2p^2 \\
&= -np^2 + np = np(1-p)
\end{aligned}$$

Example

40% of students in a college are male. 10 students are selected. Let Y be the number of male students in the sample. Find $\mathbb{E}(Y)$ and $\mathrm{Var}(Y)$.

Note: Bernoulli distribution

When we have a single trial ($n - 1$), the binomial probability function is
$\mathbb{P}(Y = 1) = p$
$\mathbb{P}(Y = 0) = 1 - p$

This distribution is usually called the Bernoulli distribution.

Binomial as a sum of independent Bernoulli r.v.

If X_1, X_2, ..., X_n are independent Bernoulli r.v with parameter p, then

$$X_1 + \ldots + X_n \sim \text{Binomial}(n, p)$$

This gives another way to compute the expectation and variance of a binomial r.v.

Quiz

Let $X \sim \text{Binom}(n, p)$ and $Y \sim \text{Binom}(m, p)$ be independent. Then X + Y follows:

(A) $\text{Binom}(n + m, p)$
(B) $\text{Binom}(nm, p)$
(C) $\text{Binom}(n + m, 2p)$
(D) other

Quiz

Let $X \sim \text{Binom}(n, p_1)$ and $Y \sim \text{Binom}(n, p_2)$ be independent.
Then X + Y follows:

(A) $\text{Binom}(n, p_1 + p_2)$
(B) $\text{Binom}(2n, p_1 + p_2)$
(C) $\text{Binom}(n, p_1 p_2)$
(D) other

The Geometric Distribution

- Consider an experiment with a series of identical and independent trials, each resulting in either a Success or a Failure.
- This is similar to binomial experiment, except there is not a fixed number of trials. Rather, the series concludes after the first success.
- The r.v. of interest, Y, is the number of the trial on which the first success occurs.

The Geometric Distribution

- Consider an experiment with a series of identical and independent trials, each resulting in either a Success or a Failure.
- This is similar to binomial experiment, except there is not a fixed number of trials. Rather, the series concludes after the first success.
- The r.v. of interest, Y, is the number of the trial on which the first success occurs.

The Geometric Distribution

- Consider an experiment with a series of identical and independent trials, each resulting in either a Success or a Failure.
- This is similar to binomial experiment, except there is not a fixed number of trials. Rather, the series concludes after the first success.
- The r.v. of interest, Y, is the number of the trial on which the first success occurs.

Pmf of the Geometric Distribution

The sample space consists of: S, FS, FFS, FFFS, ...

If $\mathbb{P}(Success) = p$ and $\mathbb{P}(Failure) = q = 1 - p$, then
$\mathbb{P}(Y = 1) = p$, $\mathbb{P}(Y = 2) = qp$, $\mathbb{P}(Y = 3) = q^2 p$.

In general $\mathbb{P}(Y = k) = q^{k-1} p$.

This is the pmf of the geometric r.v.

(This book uses a slightly non-standard definition. Often, the geometric r.v. is defined as the number of trials before the first success.)

The mean and the variance of the geometric distribution

Theorem If Y is a geometric r.v., $Y \sim Geom(p)$, then

1. $\mathbb{E}(Y) = \frac{1}{p}$
2. $\mathrm{Var}(Y) = \frac{1-p}{p^2}$

Proof: (1)
$\mathbb{E}(Y) = \sum_{y=1}^{\infty} yq^{y-1}p$
Recall the formula for the geometric series: $\sum_{y=0}^{\infty} q^y = \frac{1}{1-q}$ and differentiate it over q:

$$\sum_{y=1}^{\infty} yq^{y-1} = \frac{d}{dq}\frac{1}{1-q} = \frac{1}{(1-q)^2}.$$

Hence,

$$\mathbb{E}(Y) = \frac{p}{(1-q)^2} = \frac{1}{p}$$

(2) Similar trick as for binomial r.v. Exercise. $\square$

Another way to compute the mean

We can use a recursive equation!

$$\mathbb{E}(Y) = p \cdot 1 + q \cdot (\mathbb{E}(Y) + 1)$$

Memory

Geometric r.v. is memoryless, i.e.

$$\mathbb{P}(X = n + k | X > n) = P(X = k)$$

Proof:

Quiz

Suppose that the probability of an applicant passing a driving test is 0.25 on any given attempt and that the attempts are independent. What is the expected number of attempts?

Negative binomial r.v.

- Consider an experiment with a series of identical and independent trials, each resulting in either a Success or a Failure.
- Define a negative binomial r.v. Y as the number of the trial on which the r-th success occurs.
  (If $r = 1$ then this Y is a geometric r.v.)
- For the binomial r.v. we fix the number of trials and count the number of successes obtained.
  For the negative binomial we fix the number of successes and count the number of trials needed.

Pmf of a negative binomial r.v.

What is the probability that the r-th success occured on trial k?

This means that the first $k-1$ trials contain $r-1$ successes, and the k-th trial is success.

The probability of this is:

$$\binom{k-1}{r-1} p^{r-1} q^{k-r} p = \binom{k-1}{r-1} p^r q^{k-r}$$

This is the probability function of the negative binomial distribution.

Mean and variance of the negative binomial distribution

Theorem If Y is a negative binomial distribution, $Y \sim NB(r, p)$, then

1. $\mathbb{E}(Y) = \frac{r}{p}$
2. $\mathrm{Var}(Y) = \frac{r(1-p)}{p^2}$

Proof: Negative binomial r.v. is a sum of independent geometric r.v.'s:

$$Y = X_1 + X_2 + \ldots + X_r.$$

Coupon Collector Problem

You collect coupons which comes with some items (like oatmeal). There are N different coupons, they are all equally likely, with the probability to purchase any type at any time equal to $\frac{1}{N}$. What is the expected number of coupons that you need to purchase in order to have a complete collection of coupons?

Example

Suppose 40% of employees at a firm have traces of asbestos in their lungs. The firm is asked to send 3 of such employees to a medical center for further testing.

Find the probability that exactly 10 employees must be checked to find 3 with asbestos traces.

What is the expected number of employees that must be checked?

If Y is the number of employees that must be checked, what is $\mathrm{Var}(Y)$ and σ?

Calculation of negative binomial and geometric variables in R

Suppose that the probability of an applicant passing a driving test is 0.25 on any given attempt and that the attempts are independent. What is the probability that his initial pass is on his fourth try?

What is $\mathbb{P}(Y = 4)$?
dgeom(4-1, prob = 0.25)
(Here 4-1 is because the geometric distribution is defined slightly differently in R)

Suppose 40% of employees at a firm have traces of asbestos in their lungs. The firm is asked to send 3 of such employees to a medical center for further testing.
Find the probability that at most 10 employees will be checked to find 3 with asbestos traces.

What is $\mathbb{P}(Y \leq 10)$?
pnbinom(10-3, size = 3, prob = 0.40)

Hypergeometric distribution: example

Example Suppose 2% of items produced from an assembly line are defective. If we sample 10 items, what is the probability that 2 or more are defective?
Here the count follows the binomial distribution.

Suppose now that we sample 10 items from a small collection, like 20 items, and count the number of defectives. The resulting random variable is not binomial.

Why not?

Hypergeometric distribution: definition

Let an urn contain N balls and r of them are red. Suppose we take a sample of n balls and let Y be the number of red balls in the sample. What is the probability that $Y = y$?

$$p(y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}$$

Proof: There are $\binom{N}{n}$ different choices of n balls out of N.
So the probability of each sample point is $\frac{1}{\binom{N}{n}}$.

If the sample contains y red balls then they could be chosen in $\binom{r}{y}$ different ways. At the same time the $n - y$ black balls could be chosen in $\binom{N-r}{n-y}$ different ways.

So the total number of sample points with y red balls is $\binom{r}{y}\binom{N-r}{n-y}$.

By the sample point method, $\mathbb{P}(Y = y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}$ $\square$

Mean and Variance of the hypergeometric distribution

Theorem: If $Y \sim hyper(r, N, n)$, then

1. $\mathbb{E}(Y) = n\left(\frac{r}{N}\right)$
2. $\mathrm{Var}(Y) = n\left(\frac{r}{N}\right)\left(1 - \frac{r}{N}\right)\left(\frac{N-n}{N-1}\right)$

Proof of 1.: By induction, using

$$\mathbb{E}Y_{N,r,n} = \frac{r}{N}(\mathbb{E}Y_{N-1,r-1,n-1} + 1) + \frac{N-r}{N}\mathbb{E}Y_{N-1,r,n-1}.$$

Note the remarkable resemblance with the mean and the variance of the binomial distribution $Bin(p, n)$ if we set $p = r/N$.

The only difference is that the variance is multiplied by $\left(\frac{N-n}{N-1}\right)$.

This factor is called “finite population adjustment”.

Example

From a set of 20 potential jurors (8 African-American and 12 white) 6 jurors were selected. If the jury selection was random, what is the probability of one or fewer African Americans on the jury?

What is the expected number of African-American on the jury? What is the standard deviation?

Quiz

An urn contains 20 marbles, of which 10 are green, 5 are blue, and 5 are red. 4 marbles are to be drawn from the urn, one at a time without replacement. Let Y be the total number of green balls drawn in the sample.

$\mathrm{Var}(Y) = ?$

The Poisson distribution

Consider a r.v. Y that counts the number of occurrences of some phenomenon during some fixed unit of time (or space).

Examples:

- Y = number of phone calls received per day.
- Y= number of accidents per week at an intersection.
- Y = number of spots per square inch of an orange.

Assume the mean number of occurrences is fixed at λ.

Divide this interval of time (or space) into a large number n of subintervals. Let the probability of an occurrence in each subinterval is p and assume that occurences in different intervals are independent.

If we ignore the probability that two or more occurrences happen in one subinterval, then we have a binomial experiment.

Let $n \to \infty$, $p \to 0$, so that $np \to \lambda$.

The Poisson distribution:derivation

Then p(k)=

$$
\begin{aligned}
\binom{n}{k} p^k (1-p)^{n-k} &= \frac{n(n-1)\ldots(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k} \\
&= \frac{\lambda^k}{k!}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}\left(1-\frac{1}{n}\right)\ldots\left(1-\frac{k-1}{n}\right) \\
&\to \frac{\lambda^k}{k!}e^{-\lambda}
\end{aligned}
$$

By definition, the probability function of the Poisson r.v. is $p(k) = \frac{\lambda^k}{k!}e^{-\lambda}$, where $k = 0, 1, 2, \ldots$

Quiz

The daily probability of no accidents on a given piece of a highway is e^{-2}. The number of accidents follows the Poisson distribution.

What is the expected number of accidents per day?

The mean and variance of the Poisson distribution

Theorem: If $Y \sim Pois(\lambda)$, then

1. $\mathbb{E}(Y) = \lambda$
2. $\mathrm{Var}(Y) = \lambda$

Intuitively: Poisson r.v. is a limit of binomial r.v.'s, hence its expectation and variance are also limits: $np \to \lambda$, $np(1-p) \to \lambda$.

Proof: (1)

$$\begin{aligned} \mathbb{E}(Y) &= \sum_{y=0}^{\infty} y \frac{\lambda^y}{y!} e^{-\lambda} \\ &= \lambda e^{-\lambda} \sum_{y=1}^{\infty} \frac{\lambda^{y-1}}{(y-1)!} \\ &= \lambda. \end{aligned}$$

(2) $\mathbb{E}[Y(Y-1)] = \lambda^2$, $\mathbb{E}(Y^2) = \lambda^2 + \lambda$,
$\mathrm{Var}(Y) = \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$

Example

Suppose the number of accidents per month at an industrial plant has a Poisson distribution with mean 2.6. Find the probability that there will be 4 accidents in the next month.

Examples

Suppose the number of accidents per month at an industrial plant has a Poisson distribution with mean 2.6. Find the probability of having between 3 and 6 accidents in the next month.

The cumulative Poisson probabilities can be evaluated either using Table 3 in Appendix 3 or using R: $ppois(y, \lambda)$

$$\mathbb{P}(Y \leq 6) - \mathbb{P}(Y \leq 2) =$$

What is the probability of 10 accidents in the next half-year?

Quiz

Suppose the number of accidents per month at an industrial plant has a Poisson distribution with mean 2.6.

What is the expected number of accidents in the next half-year?

Poisson Process

Poisson r.v.'s usually come not alone but in collections which are called the Poisson processes.

A stochastic process is a r.v. that depends on a parameter. For example, a stock price depends on time.

For the Poisson process, the parameter is usually a region in space (or time). So each region D correspond to a Poisson r.v. X_D with the mean $\lambda(D)$. This random variable is interpreted as the number of points in this region.

It is assumed that if the regions $D_1, \ldots, D_k$ are disjoint, the corresponding r.v.'s $X_{D_1}, \ldots, X_{D_k}$ are independent.

This implies that the means of Poisson r.v.'s are additive: if D_1 and D_2 are disjoint, then $\lambda(D_1 \cup D_2) = \lambda(D_1) + \lambda(D_2)$.

Relationship with binomial distribution

If n is large, p is small, and $\lambda = np$ is somewhat small (book: ≤ 7), then the $Bin(n, p)$ probabilities are approximately equal to the $Pois(\lambda)$ probabilities.

Example: Suppose there are $10,000$ students in a college. 3% of all students are vegetarians. Select 100 students at random. Find the probability of at least 5 vegetarians.

Hypergeometric:

Binomial:

Poisson:

Quiz

A parking lot has two entrances. Cars arrive at entrance I according to a Poisson distribution at an average of 5 per hour and at entrance II according to a Poisson distribution at an average of 3 per hour. (Assume that the numbers of cars arriving at the two entrances are independent.)
Then total number of cars that arrive at parking lot is Poisson with an average

(A) 4 cars per hour

(B) 8 cars per hour

(C) 15 cars per hour

(D) Other

Quiz

Let X be a random variable whose Probability Mass Function is given as

$$P(X = k) = \frac{5^k}{e^5 k!}$$

for $k = 0, 1, \ldots$

What is the variance of this random variable?

Quiz

Consider the probability given by the expression:

$$\mathbb{P}(Y = 3) = \frac{\binom{2{,}000{,}000}{3}\binom{98{,}000{,}000}{97}}{\binom{100{,}000{,}000}{100}}$$

What is the appropriate value for parameter p in the binomial approximation for this probability?

(A) 2%
(B) 3%
(C) 4%
(D) 97%
(E) Other

Quiz

Consider the probability given by the expression:

$$\mathbb{P}(Y = 3) = \frac{\binom{2{,}000{,}000}{3}\binom{98{,}000{,}000}{97}}{\binom{100{,}000{,}000}{100}}$$

What is the appropriate value for parameter λ in the Poisson approximation for this probability?

Quiz

Let a r.v. Y have probability function $p(y) = \left(\frac{1}{2}\right)^y$ for $y = 1, 2, 3, \ldots$.
Then, the distribution of Y is

(A) Poisson

(B) Hypergeometric

(C) Geometric

(D) Binomial

(E) None of above

Moments

Defn: The k-th moment of a r.v. Y is defined as

$$\mu'_k = \mathbb{E}(Y^k)$$

Examples:

- The first moment, μ'_1 is the expected value of Y.
- We know $\mu'_2 - (\mu'_1)^2$ is the variance of Y.

Note: The expression $\mathbb{E}[(Y-\mu)^k]$ is sometimes called the k-th moment about the mean (or the k-th central moment).
Example:

- $\mathbb{E}[(Y-\mu)^3]$ is called the skeweness of Y.

Moment generating function

Defn: The moment-generating function (or mgf) of a r.v. Y is defined to be

$$m_Y(t) = \mathbb{E}(e^{tY}).$$

The mgf for Y exists if there exists some $b > 0$, such that $\mathbb{E}(e^{bY})$ exists.

Theorem If $m_Y(t)$ exists, then for any integer $k \geq 1$,

$$\left.\frac{d^k m_Y(t)}{dt^k}\right|_{t=0} = \mu'_k$$

Proof:

$$m_Y(t) = \mathbb{E}(e^{tY}) = 1 + t\mu'_1 + \frac{t^2}{2!}\mu'_2 + \frac{t^3}{3!}\mu'_3 + \dots$$

Hence, by differentiating this expression k times, we get

$$\frac{d^k m_Y(t)}{dt^k} = \mu'_k + t\mu'_{k+1} + \dots$$

Example

Example: Let a r.v. Y have probability function $p(y) = \left(\frac{1}{2}\right)^y$ for $y = 1, 2, 3, \ldots$.
Find $m_Y(t)$ and then $\mathbb{E}(Y)$.

$$\begin{aligned} m_Y(t) &= \mathbb{E}(e^{tY}) = \sum_{y=1}^{\infty} e^{ty} e^{-(log2)y} \\ &= \frac{e^{(t-log2)}}{1-e^{(t-log2)}} = -1 + \frac{1}{1-e^{(t-log2)}} \end{aligned}$$

$$\begin{aligned} \mathbb{E}(Y) = \frac{d}{dt} m_Y(t)\Big|_{t=0} &= \frac{e^{(t-log2)}}{(1-e^{(t-log2)})^2}\Big|_{t=0} \\ &= \frac{1/2}{1/4} = 2. \end{aligned}$$

Quiz

The random variable X has the moment generating function

$$m_X(t) = \frac{3}{4}e^t + \frac{1}{4}e^{t^2}$$

for all t.

Find $\mathbb{E}(X^2)$.

(A) 3/4

(B) 1

(C) 5/4

(D) 5/2

(E) other

Mgf of geometric distribution

Suppose Y has the geometric distribution with parameter p. What is its mgf?

$$m_Y(t) = pe^t + pqe^{2t} + pq^2e^{3t} + \dots$$

$$m_Y(t) = pe^t(1 + qe^t + q^2e^{2t} + \dots)$$

$$m_Y(t) = \frac{pe^t}{1 - qe^t}$$

Mgf of Bernoulli distribution

Suppose Y is a Bernoulli distribution with parameter p. What is its mgf?

$$m_Y(t) = pe^{1\cdot t} + qe^{0\cdot t} = pe^t + q$$

Reminder: Independent r.v.

Defn: Two discrete r.v.'s X and Y are called independent if the events $\{X = x\}$ and $\{Y = y\}$ are independent for all possible values of x and y.

This is equivalent to the following:
Theorem Two discrete r.v.'s X and Y are independent if and only if $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$ for all possible functions f and g.

Proof:

Theorem If two r.v.'s X and Y are independent, then

$$m_{X+Y}(t) = m_X(t)m_Y(t).$$

Proof:

Example: mgf of binomial distribution

Binomial r.v. $Y \sim Bin(n, p)$ is the sum of n independent Bernoulli r.v. with parameter p:

$$Y = X_1 + X_2 + \ldots + X_n$$

.

By the previous theorem:

$$m_Y(t) = [m_{X_1}(t)]^n = (q + pe^t)^n.$$

Example: mgf of the negative binomial distribution

Let $Y \sim \mathrm{NegBin}(r, p)$. What is its mgf?

Y is a sum of r independent r.v. with geometric distribution. Hence

$$m_Y(t) = \left[\frac{pe^t}{1 - qe^t}\right]^r.$$

Example: mgf of Poisson distribution

Let $Y \sim Pois(\lambda)$ with probability function

$$p(y) = \frac{\lambda^y}{y!} e^{-\lambda} \text{ for } y = 0, 1, 2, \ldots.$$

Find the mgf and variance of Y.

$$\begin{aligned} m_Y(t) &= \mathbb{E}(e^{tY}) = \sum_{y=0}^{\infty} e^{ty} \frac{\lambda^y}{y!} e^{-\lambda} \\ &= e^{e^t \lambda} e^{-\lambda} = e^{\lambda(e^t - 1)}. \end{aligned}$$

After differentiation we find:

$$\frac{d}{dt} m_Y(t) = e^{\lambda(e^t-1)} \lambda e^t,$$

$$\frac{d^2}{(dt)^2} m_Y(t) = e^{\lambda(e^t-1)} (\lambda e^t)^2 + e^{\lambda(e^t-1)} \lambda e^t.$$

Evaluating at $t = 0$, we get

$$\mu = \lambda, \mu_2' = \lambda^2 + \lambda, \mathrm{Var}(Y) = \lambda^2 + \lambda - (\lambda)^2 = \lambda.$$

Uniqueness of mgf

If two r.v.'s have the same mgf then they must have the same distribution.

So we can use the mgf to identify the distribution of a r.v.

Quiz

If Y is a r.v. with mgf $m_Y(t) = e^{7.1e^t - 7.1}$, then what is the distribution of Y?

(A) Poisson

(B) Hypergeometric

(C) Geometric

(D) Binomial

(E) None of above

Chebyshev's theorem

Theorem: Let X be a r.v. with mean μ and variance σ^2. Then for every $k > 0$, we have

$$\mathbb{P}(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Example: Let X have mean 20 and standard deviation 2. What can be said about the probability that a new realization of X will be between 16 and 24?

Chebyshevs theorem applies with k=2. Hence,

$$\mathbb{P}(|X - 20| \geq 4) \leq \frac{1}{2^2} = \frac{1}{4}.$$

$$\mathbb{P}(16 < X < 24) \geq 1 - \frac{1}{2^2} = 3/4.$$

Quiz

Let $X \sim \mathrm{Binom}(n = 4, p = 1/2)$.

What is the upper bound on $\mathbb{P}(X \geq 4)$ given by Chebyshev inequality?

(A) 1/2
(B) 1/4
(C) 1/8
(D) 1/16
(E) Other

Quiz

Let $X \sim \text{Binom}(n = 4, p = 1/2)$.

What is the exact value of $\mathbb{P}(X \geq 4)$?

(A) 1/2
(B) 1/4
(C) 1/8
(D) 1/16
(E) Other