P2 - Discrete Random Variables

STAT 5870 (Engineering)
Iowa State University

September 5, 2024

Random variables

If Ω is the sample space of an experiment, a random variable X is a function $X(\omega) : \Omega \mapsto \mathbb{R}$.

Idea: If the value of a numerical variable depends on the outcome of an experiment, we call the variable a *random variable*.

Examples of random variables from rolling two 6-sided dice:

- Sum of the two dice
- Indicator of the sum being greater than 5

We will use an upper case Roman letter (late in the alphabet) to indicate a random variable and a lower case Roman letter to indicate a realized value of the random variable.

8 bit example

Suppose, 8 bits are sent through a communication channel. Each bit has a certain probability to be received incorrectly. We are interested in the number of bits that are received incorrectly.

- Let X be the number of incorrect bits received.
- The possible values for X are $\{0, 1, 2, 3, 4, 5, 6, 7, 8\}$.
- Example events:
  - No incorrect bits received: $\{X = 0\}$.
  - At least one incorrect bit received: $\{X \geq 1\}$.
  - Exactly two incorrect bits received: $\{X = 2\}$.
  - Between two and seven (inclusive) incorrect bits received: $\{2 \leq X \leq 7\}$.

Range/image of random variables

The range (or image) of a random variable X is defined as

$$Range(X) := \{x : x = X(\omega) \text{ for some } \omega \in \Omega\}$$

If the range is finite or countably infinite, we have a discrete random variable. If the range is uncountably infinite, we have a continuous random variable.
Examples:

- Put a hard drive into service, measure $Y =$ "time until the first major failure" and thus $Range(Y) = (0, \infty)$. Range of Y is an interval (uncountable range), so Y is a continuous random variable.
- Communication channel: $X =$ "# of incorrectly received bits out of 8 bits sent" with $Range(X) = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$.
  Range of X is a finite set, so X is a discrete random variable.
- Communication channel: $Z =$ "# of incorrectly received bits in 10 minutes" with $Range(Z) = \{0, 1, \ldots\}$.
  Range of Z is a countably infinite set, so Z is a discrete random variable.

Distribution

The collection of all the probabilities related to X is the distribution of X.

For a discrete random variable, the function

$$p_X(x) = P(X = x)$$

is the probability mass function (pmf) and the cumulative distribution function (cdf) is

$$F_X(x) = P(X \leq x) = \sum_{y \leq x} p_X(y).$$

The set of non-zero probability values of X is called
the support of the distribution f.
This is the same as the range of X.

Examples

A probability mass function is valid if it defines a valid set of probabilities, i.e. they obey Kolmogorov's axioms.

Which of the following functions are a valid probability mass functions?

x	-3	-1	0	5	7
$p_X(x)$	0.1	0.45	0.15	0.25	0.05

y	-1	0	1.5	3	4.5
$p_Y(y)$	0.1	0.45	0.25	-0.05	0.25

z	0	1	3	5	7
$p_Z(z)$	0.22	0.18	0.24	0.17	0.18

Rolling a fair 6-sided die

Let Y be the number of pips on the upturned face of a die. The support of Y is $\{1, 2, 3, 4, 5, 6\}$. If we believe the die has equal probability for each face, then image, pmf, and cdf for Y are

y	1	2	3	4	5	6
$p_Y(y) = P(Y = y)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$F_Y(y) = P(Y \leq y)$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{4}{6}$	$\frac{5}{6}$	$\frac{6}{6}$

Dragonwood

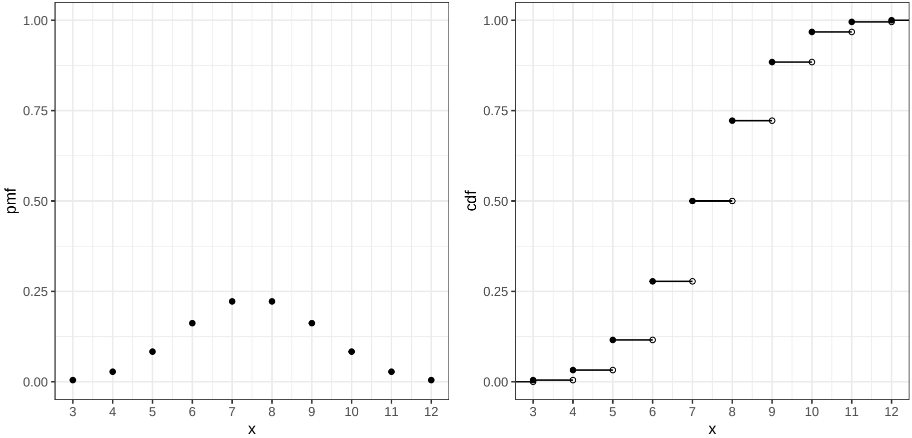
Dragonwood has 6-sided dice with the following # on the 6 sides: $\{1, 2, 2, 3, 3, 4\}$.

What is the support, pmf, and cdf for the sum of the upturned numbers when rolling 3 Dragonwood dice?

```
# Three dice
die   = c(1,2,2,3,3,4)
rolls = expand.grid(die1 = die, die2 = die, die3 = die)
sum   = rowSums(rolls); tsum = table(sum)
dragonwood3 = data.frame(x   = as.numeric(names(tsum)),
                         pmf = as.numeric(table(sum)/length(sum))) %>%
        mutate(cdf = cumsum(pmf))
```

x	$P(X = x)$	$P(X \leq x)$
3	0.005	0.005
4	0.028	0.032
5	0.083	0.116
6	0.162	0.278
7	0.222	0.500
8	0.222	0.722
9	0.162	0.884
10	0.083	0.968
11	0.028	0.995
12	0.005	1.000

Dragonwood - pmf and cdf



Properties of pmf and cdf

Properties of probability mass function $p_X(x) = P(X = x)$:

- $0 \leq p_X(x) \leq 1$ for all $x \in \mathbb{R}$.
- $\sum_{x \in S} p_X(x) = 1$ where S is the support.

Properties of cumulative distribution function $F_X(x)$:

- $0 \leq F_X(x) \leq 1$ for all $x \in \mathbb{R}$
- F_X is nondecreasing, (i.e. if $x_1 \leq x_2$ then $F_X(x_1) \leq F_X(x_2)$.)
- $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to \infty} F_X(x) = 1$.
- $F_X(x)$ is right continuous with respect to x

Dragonwood (cont.)

In Dragonwood, you capture monsters by rolling a sum equal to or greater than its defense. Suppose you can roll 3 dice and the following monsters are available to be captured:

- Spooky Spiders worth 1 victory point with a defense of 3.
- Hungry Bear worth 3 victory points with a defense of 7.
- Grumpy Troll worth 4 victory points with a defense of 9.

Which monster should you attack?

Dragonwood (cont.)

Calculate the probability by computing one minus the cdf evaluated at "defense minus 1". Let X be the sum of the number on 3 Dragonwood dice. Then

- $P(X \geq 3) = 1 - P(X \leq 2) = 1$
- $P(X \geq 7) = 1 - P(X \leq 6) = 0.722.$
- $P(X \geq 9) = 1 - P(X \leq 8) = 0.278.$

If we multiply the probability by the number of victory points,
then we have the "expected points":

- $1 \times P(X \geq 3) = 1$
- $3 \times P(X \geq 7) = 2.17.$
- $4 \times P(X \geq 9) = 1.11.$

Expectation

Let X be a random variable and h be some function. The expected value of a function of a (discrete) random variable is

$$E[h(X)] = \sum_i h(x_i) \cdot p_X(x_i).$$

Intuition: Expected values are *weighted averages* of the possible values weighted by their probability.

If $h(x) = x$, then

$$E[X] = \sum_i x_i \cdot p_X(x_i)$$

and we call this the expectation of X and commonly use the symbol μ for the expectation.

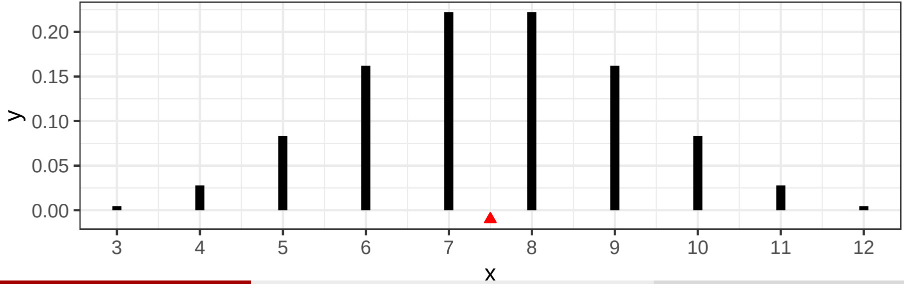
Dragonwood (cont.)

What is the expectation of the sum of 3 Dragonwood dice?

```
expectation = with(dragonwood3, sum(x*pmf))
expectation
```

```
[1] 7.5
```

The expectation can be thought of as the center of mass if we place mass $p_X(x)$ at corresponding points x.



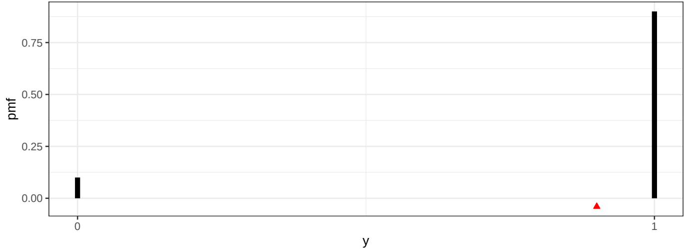
Biased coin

Suppose we have a biased coin represented by the following pmf:

y	0	1
$p_Y(y)$	$1-p$	p

What is the expected value?

If $p = 0.9$,



Properties of expectations

Let X and Y be random variables and a, b, and c be constants. Then

$$E[aX + bY + c] \quad = aE[X] + bE[Y] + c.$$

In particular

- $E[X + Y] = E[X] + E[Y]$,
- $E[aX] = aE[X]$, and
- $E[c] = c$.

Dragonwood (cont.)

Enhancement cards in Dragonwood allow you to improve your rolls. Here are two enhancement cards:

- *Cloak of Darkness* adds 2 points to all capture attempts and
- *Friendly Bunny* allows you (once) to roll an extra die.

What is the expected attack roll total if you had 3 Dragonwood dice, the Cloak of Darkness, and are using the Friendly Bunny?

Let

- X be the sum of 3 Dragonwood dice (we know $E[X] = 7.5$),
- Y be the sum of 1 Dragonwood die which has $E[Y] = 2.5$.

Then the attack roll total is $X + Y + 2$ and the *expected* attack roll total is

$$E[X + Y + 2] = E[X] + E[Y] + 2 = 7.5 + 2.5 + 2 = 12.$$

Variance

The variance of a random variable is defined as the expected squared deviation from the mean. For discrete random variables, variance is

$$Var[X] = E[(X-\mu)^2] = \sum_i (x_i - \mu)^2 \cdot p_X(x_i)$$

where $\mu = E[X]$. The symbol σ^2 is commonly used for the variance.
The variance is analogous to moment of intertia in classical mechanics.

The standard deviation (sd) is the positive square root of the variance:

$$SD[X] = \sqrt{Var[X]}.$$

The symbol σ is commonly used for sd.

Properties of variance

Two discrete random variables X and Y are independent if

$$p_{X,Y}(x, y) = p_X(x)p_Y(y).$$

If X and Y are independent, and a, b, and c are constants, then

$$Var[aX + bY + c] = a^2Var[X] + b^2Var[Y].$$

Special cases:

- $Var[c] = 0$
- $Var[aX] = a^2Var[X]$
- $Var[X + Y] = Var[X] + Var[Y]$
  (if X and Y are independent)

Dragonwood (cont.)

What is the variance for the sum of the 3 Dragonwood dice?

```
variance = with(dragonwood3, sum((x-expectation)^2*pmf))
variance

[1] 2.75
```

What is the standard deviation for the sum of the pips on 3 Dragonwood dice?

```
sqrt(variance)

[1] 1.658312
```

Biased coin

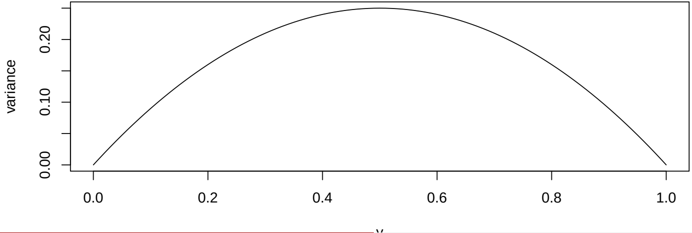
Suppose we have a biased coin represented by the following pmf:

y	0	1
$p_Y(y)$	$1-p$	p

What is the variance?

1. $E[Y] = p$
2. $Var[y] = (0-p)^2(1-p) + (1-p)^2 \times p = p - p^2 = p(1-p)$

When is this variance maximized?



Special discrete distributions

- Bernoulli
- Binomial
- Poisson

Note: The range is always finite or countable.

Bernoulli random variables

A Bernoulli experiment has only two outcomes: success/failure.

Let

- $X = 1$ represent success and
- $X = 0$ represent failure.

The probability mass function $p_X(x)$ is

$$p_X(0) = 1 - p \quad p_X(1) = p.$$

We use the notation $X \sim Ber(p)$ to denote a random variable X that follows a Bernoulli distribution with success probability p, i.e. $P(X = 1) = p$.

Bernoulli experiment examples

- Toss a coin: $\Omega = \{Heads, Tails\}$
- Throw a fair die and ask if the face value is a six:
  $\Omega = \{\text{face value is a six}, \text{face value is not a six}\}$
- Send a message through a network and record whether or not it is received:
  $\Omega = \{\text{successful transmission},\ \text{unsuccessful transmission}\}$
- Draw a part from an assembly line and record whether or not it is defective:
  $\Omega = \{\text{defective},\ \text{good}\}$
- Response to the question
  "Are you in favor of an increased in property tax
  xto pay for a new high school?":
  $\Omega = \{\text{yes}, \text{no}\}$

Bernoulli random variable (cont.)

The cdf of the Bernoulli random variable is

$$F_X(x) = P(X \leq x) = \begin{cases} 0 & x < 0 \\ 1-p & 0 \leq x < 1 \\ 1 & 1 \leq x \end{cases}$$

The expected value is

$$E[X] = \sum_x p_X(x) = 0 \cdot (1-p) + 1 \cdot p = p.$$

The variance is

$$\begin{aligned} Var[X] &= \sum_x (x - E[X])^2 p_X(x) \\ &= (0-p)^2 \cdot (1-p) + (1-p)^2 \cdot p \\ &= p(1-p). \end{aligned}$$

Sequence of Bernoulli experiments

An experiment consisting of n independent and identically distributed Bernoulli experiments.

Examples:

- Toss a coin n times and record the nubmer of heads.
- Send 23 identical messages through the network independently and record the number successfully received.
- Draw 5 cards from a standard deck with replacement (and reshuffling) and record whether or not the card is a king.

Independent and identically distributed

Let X_i represent the i^{th} Bernoulli experiment.

Independence means

$$p_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \prod_{i=1}^{n} p_{X_i}(x_i),$$

i.e. the joint probability is the product of the individual probabilities.

Identically distributed (for Bernoulli random variables) means

$$P(X_i = 1) = p \quad \forall\, i,$$

and more generally, the distribution is the same for all the random variables.

- iid: independent and identically distributed
- ind: independent

Sequences of Bernoulli experiments

Let X_i denote the outcome of the i^{th} Bernoulli experiment. We use the notation

$$X_i \stackrel{iid}{\sim} Ber(p), \quad \text{for } i = 1, \ldots, n$$

to indicate a sequence of n independent and identically distributed Bernoulli experiments.

We could write this equivalently as

$$X_i \stackrel{ind}{\sim} Ber(p), \quad \text{for } i = 1, \ldots, n$$

but this is different than

$$X_i \stackrel{ind}{\sim} Ber(p_i), \quad \text{for } i = 1, \ldots, n$$

as the latter has a different success probability for each experiment.

Binomial random variable

Suppose we perform a sequence of n iid Bernoulli experiments and only record the number of successes, i.e.

$$Y = \sum_{i=1}^{n} X_i.$$

Then we use the notation $Y \sim Bin(n, p)$ to indicate a binomial random variable with

- n attempts and
- probability of success p.

Binomial probability mass function

We need to obtain

$$p_Y(y) = P(Y = y) \quad \forall\, y \in \Omega = \{0, 1, 2, \ldots, n\}.$$

The probability of obtaining a particular sequence of y success and $n - y$ failures is

$$p^y(1-p)^{n-y}$$

since the experiments are iid with success probability p. But there are

$$\binom{n}{y} = \frac{n!}{y!(n-y)!}$$

ways of obtaining a sequence of y success and $n - y$ failures. Thus, the binomial pmf is

$$p_Y(y) = P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y}.$$

Properties of binomial random variables

The expected value is

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

The variance is

$$Var[Y] = \sum_{i=1}^{n} Var[X_i] = np(1-p)$$

since the X_i are independent.

The cumulative distribution function is

$$F_Y(y) = P(Y \le y) = \sum_{x=0}^{\lfloor y \rfloor} \binom{n}{x} p^x (1-p)^{n-x}.$$

Component failure rate

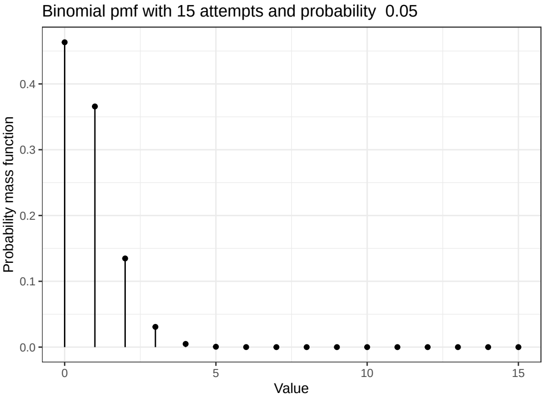
Suppose a box contains 15 components that each have a failure rate of 5%.

What is the probability that

1. exactly two out of the fifteen components are defective?
2. at most two components are defective?
3. more than three components are defective?
4. more than 1 but less than 4 are defective?

Binomial pmf

Let Y be the number of defective components and assume $Y \sim Bin(15, 0.05)$.



Component failure rate - solutions

Let Y be the number of defective components and assume $Y \sim Bin(15, 0.05)$.

1. $P(Y = 2) = \binom{15}{2}(0.05)^2(1-0.05)^{15-2}$
2. $P(Y \leq 2) = \sum_{x=0}^{2} \binom{15}{x}(0.05)^x(1-0.05)^{15-x}$
3. $P(Y > 3) = 1 - P(Y \leq 3) = 1 - \sum_{x=0}^{3} \binom{15}{x}(0.05)^x(1-0.05)^{15-x}$
4. $P(1 < Y < 4) = \sum_{x=2}^{3} \binom{15}{x}(0.05)^x(1-0.05)^{15-x}$

Component failure rate - solutions in R

```
n <- 15
p <- 0.05
choose(15,2)

[1] 105

dbinom(2,n,p)           # P(Y=2)

[1] 0.1347523

pbinom(2,n,p)           # P(Y<=2)

[1] 0.9637998

1-pbinom(3,n,p)         # P(Y>3)

[1] 0.005467259

sum(dbinom(c(2,3),n,p)) # P(1<Y<4) = P(Y=2)+P(Y=3)

[1] 0.1654853
```

Poisson experiments

Many experiments can be thought of as "how many *rare* events will occur in a certain amount of time or space?" For example,

- # of alpha particles emitted from a polonium bar in an 8 minute period
- # of flaws on a standard size piece of manufactured product, e.g., 100m coaxial cable, 100 sq.meter plastic sheeting
- # of hits on a web page in a 24h period

Poisson random variable

A Poisson random variable has pmf

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!} \quad \text{for } x = 0, 1, 2, 3, \ldots$$

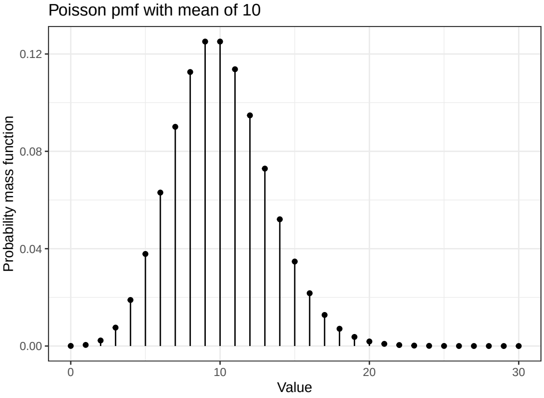
where λ is called the rate parameter.

We write $X \sim Po(\lambda)$ to represent this random variable. We can show that

$$E[X] = Var[X] = \lambda.$$

Poisson probability mass function

Customers of an internet service provider initiate new accounts at the average rate of 10 accounts per day. What is the probability that more than 8 new accounts will be initiated today?



Poisson probability

Customers of an internet service provider initiate new accounts at the average rate of 10 accounts per day. What is the probability that more than 8 new accounts will be initiated today?

Let X be the number of accounts initiated today. Assume $X \sim Po(10)$.

$$P(X > 8) = 1 - P(X \leq 8) = 1 - \sum_{x=0}^{8} \frac{\lambda^x e^{-\lambda}}{x!} \approx 1 - 0.333 = 0.667$$

In R,

```
# Using pmf
1-sum(dpois(0:8, lambda=10))

[1] 0.6671803

# Using cdf
1-ppois(8, lambda=10)

[1] 0.6671803
```

Sum of Poisson random variables

Let $X_i \stackrel{ind}{\sim} Po(\lambda_i)$ for $i = 1, \ldots, n$. Then

$$Y = \sum_{i=1}^{n} X_i \sim Po\left(\sum_{i=1}^{n} \lambda_i\right).$$

Let $X_i \stackrel{iid}{\sim} Po(\lambda)$ for $i = 1, \ldots, n$. Then

$$Y = \sum_{i=1}^{n} X_i \sim Po\left(n\lambda\right).$$

Poisson random variable - example

Customers of an internet service provider initiate new accounts at the average rate of 10 accounts per day. What is the probability that more than 16 new accounts will be initiated in the next two days?

Since the rate is 10/day, then for two days we expect, on average, to have 20. Let Y be he number initiated in a two-day period and assume $Y \sim Po(20)$. Then

$$\begin{aligned} P(Y > 16) &= 1 - P(Y \leq 16) \\ &= 1 - \sum_{x=0}^{16} \frac{\lambda^x e^{-\lambda}}{x!} \\ &= 1 - 0.221 = 0.779. \end{aligned}$$

In R,

```r
# Using pmf
1-sum(dpois(0:16, lambda=20))

[1] 0.7789258

# Using cdf
1-ppois(16, lambda=20)

[1] 0.7789258
```

Manufacturing example

A manufacturer produces 100 chips per day and, on average, 1% of these chips are defective. What is the probability that no defectives are found in a particular day?

Let X represent the number of defectives and assume $X \sim Bin(100, 0.01)$. Then

$$P(X = 0) = \binom{100}{0}(0.01)^0(1 - 0.01)^{100} \approx 0.366.$$

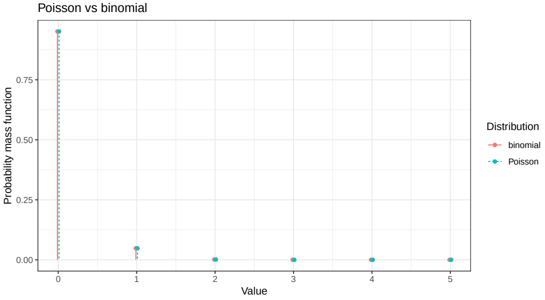
Alternatively, let Y represent the number of defectives and assume $Y \sim Po(100 \times 0.01)$. Then

$$P(Y = 0) = \frac{1^0 e^{-1}}{0!} \approx 0.368.$$

Poisson approximation to the binomial

Suppose we have $X \sim Bin(n, p)$ with n large (say ≥ 20) and p small (say ≤ 0.05). We can approximate X by $Y \sim Po(np)$ because for large n and small p

$$\binom{n}{k} p^k (1-p)^{n-k} \approx e^{-np} \frac{(np)^k}{k!}.$$



Example

Imagine you are supposed to proofread a paper. Let us assume that there are on average 2 typos on a page and a page has 1000 words. This gives a probability of 0.002 for each word to contain a typo. What is the probability the page has no typos?

Let X represent the number of typos on the page and assume $X \sim Bin(1000, 0.002)$. $P(X = 0)$ using R is

```
n = 1000; p = 0.002
dbinom(0, size=n, prob=p)

[1] 0.1350645
```

Alternatively, let Y represent the number of defectives and assume $Y \sim Po(1000 \times 0.002)$. $P(Y = 0)$ using R is

```
dpois(0, lambda = n*p)

[1] 0.1353353
```

Summary

- General discrete random variables
  - Probability mass function (pmf)
  - Cumulative distribution function (cdf)
  - Expected value
  - Variance
  - Standard deviation
- Specific discrete random variables
  - Bernoulli
  - Binomial
  - Poisson