Probability 1

Lecture 4: Discrete Random Variables - Part 2

Dr. Daniel Flores-Agreda,

(based on the notes of Prof. Davide La Vecchia)

Spring Semester 2021

Objectives

- Describe some important discrete distributions
  - PMF, PDF, Expectation, Variance.
  - Illustrate case uses.

Outline

Some important Discrete Distributions

Remark

Their main common characteristic is that the probability $P(\{X = x_i\})$ is given by an appropriate mathematical formula: i.e.

$$p_i = P(\{X = x_i\}) = h(x_i)$$

for a suitably specified function $h(\cdot)$.

Discrete Uniform

Discrete Uniform

Definition

We say X has a **discrete uniform distribution** when

- X can take the values $x = 0, 1, 2, ..., k$ (for some specified finite value $k \in \mathbb{N}$)
- The probability that $X = x$ is $1/(k+1)$ is given by the **PMF**:

$$P(\{X = x\}) = \frac{1}{(k+1)}.$$

Definition

We say X has a **discrete uniform distribution** when

- X can take the values $x = 0, 1, 2, ..., k$ (for some specified finite value $k \in \mathbb{N}$)
- The probability that $X = x$ is $1/(k+1)$ is given by the **PMF**:

$$P(\{X = x\}) = \frac{1}{(k+1)}.$$

- The probability **distribution** is given by

x_i	$P(\{X = x_i\})$
0	$\frac{1}{(k+1)}$
1	$\frac{1}{(k+1)}$
$\vdots$	$\vdots$
k	$\frac{1}{(k+1)}$
Total	1

- The **Expectation** of X is

$$
\begin{aligned}
E\left[X\right] &= x_1 p_1 + \ldots + x_k p_k \\
&= 0 \cdot \frac{1}{(k+1)} + 1 \cdot \frac{1}{(k+1)} + \cdots + k \cdot \frac{1}{(k+1)} \\
&= \frac{1}{(k+1)} \cdot (0 + 1 + \cdots + k) \\
&= \frac{1}{(k+1)} \cdot \frac{k\,(k+1)}{2} \\
&= \frac{k}{2}.
\end{aligned}
$$

- The **Expectation** of X is

$$
\begin{aligned}
E[X] &= x_1 p_1 + \ldots + x_k p_k \\
&= 0 \cdot \frac{1}{(k+1)} + 1 \cdot \frac{1}{(k+1)} + \cdots + k \cdot \frac{1}{(k+1)} \\
&= \frac{1}{(k+1)} \cdot (0 + 1 + \cdots + k) \\
&= \frac{1}{(k+1)} \cdot \frac{k(k+1)}{2} \\
&= \frac{k}{2}.
\end{aligned}
$$

Example

Consider $k = 6$, then X can take on one of the seven distinct values $x = 0, 1, 2, 3, 4, 5, 6$, each with equal probability $\frac{1}{7}$, and $E(X) = 3$ (which is one of the possible outcomes).

- The variance of X is:

$$
\begin{aligned}
Var(X) &= \left(0-\frac{k}{2}\right)^2 \cdot \frac{1}{(k+1)} + \left(1-\frac{k}{2}\right)^2 \cdot \frac{1}{(k+1)} + \\
&\quad \cdots + \left(k-\frac{k}{2}\right)^2 \cdot \frac{1}{(k+1)} \\
&= \frac{1}{(k+1)} \cdot \left\{ \left(0-\frac{k}{2}\right)^2 + \left(1-\frac{k}{2}\right)^2 + \cdots + \left(k-\frac{k}{2}\right)^2 \right\} \\
&= \frac{1}{(k+1)} \cdot \frac{k(k+1)(k+2)}{12} \\
&= \frac{k(k+2)}{12}
\end{aligned}
$$

- The variance of X is:

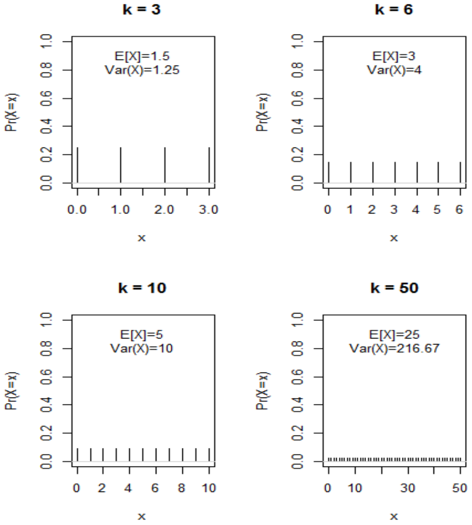
$$
\begin{aligned}
Var\left(X\right) &= \left(0-\frac{k}{2}\right)^{2}\cdot\frac{1}{\left(k+1\right)}+\left(1-\frac{k}{2}\right)^{2}\cdot\frac{1}{\left(k+1\right)}+ \\
&\quad \cdots+\left(k-\frac{k}{2}\right)^{2}\cdot\frac{1}{\left(k+1\right)} \\
&= \frac{1}{\left(k+1\right)}\cdot\left\{\left(0-\frac{k}{2}\right)^{2}+\left(1-\frac{k}{2}\right)^{2}+\cdots+\left(k-\frac{k}{2}\right)^{2}\right\} \\
&= \frac{1}{\left(k+1\right)}\cdot\frac{k\left(k+1\right)\left(k+2\right)}{12} \\
&= \frac{k\left(k+2\right)}{12}
\end{aligned}
$$

Example

When $k=6$, then $Var(X)=4$ and $s.d(X)=\sqrt{4}=2$.

Discrete Uniform

Illustrations



Example

- The outcome of the roll of a die can be modeled as a Discrete Uniform RV.
- Notice however, that $X = 0$ is not allowed in this specific example.

Let X the corresponding random variable and $\{x_1, x_2, ..., x_6\}$ its realizations:

The possible outcomes are:

$$\{1, 2, 3, 4, 5, 6\}$$

each having probability $\frac{1}{6}$.

Moreover,

$$E(X) = (1 + 2 + 3 + 4 + 5 + 6) \cdot \frac{1}{6} = 3.5,$$

which is not one of the possible outcomes.

Bernoulli Trials

Bernoulli Trials

A Bernoulli trial represents the most primitive form of all random variables. It derives from a random experiment having only two possible mutually exclusive outcomes (often labelled Success and Failure) where:

- Success occurs with probability p
- Failure occurs with probability $1 - p$.

Bernoulli Trials

A Bernoulli trial represents the most primitive form of all random variables. It derives from a random experiment having only two possible mutually exclusive outcomes (often labelled Success and Failure) where:

- Success occurs with probability p
- Failure occurs with probability $1 - p$.

Definition

Bernoulli trial has the following probability distribution

x_i	$P(\{X = x_i\})$
1	p
0	$1 - p$

Which can be written as a PMF:

$$P(\{X = x\}) = p^x \left(1 - p\right)^{1-x}, \quad \text{for } x \in \{0, 1\}$$

Remark

Often, we set $X = 1$ if Success occurs, and $X = 0$ if Failure occurs.

Example

Coin tossing to get "heads": we define the random variable

x_i	$P(\{X = x_i\})$
1	p
0	$1 - p$

and say that $X = 1$ if H and $X = 0$ if T.

- Mean:

$$
\begin{aligned}
E[X] &= 1 \cdot p + 0 \cdot (1-p) \\
&= p
\end{aligned}
$$

- Mean:

$$\begin{aligned} E\left[X\right] &= 1 \cdot p + 0 \cdot (1-p) \\ &= p \end{aligned}$$

- Variance:

$$\begin{aligned} Var\left(X\right) &= (1-p)^2 \cdot p + (0-p)^2 \cdot (1-p) \\ &= p\left(1-p\right). \end{aligned}$$

Binomial

Binomial

Definition

We could define an event that derives from carrying a **series of n independent Bernoulli trials**, i.e.

1. Only two mutually exclusive outcomes are possible in each trial: *Success* (S) and *Failure* (F).
2. Each of the of n trials constitute **independent events**.
3. The probability of success in each trial p **is constant** from trial to trial.

Let's define X **as the number of successes** occurring in all these n trials.

The event $X = x$ is associated with the sequence

$$\{\underbrace{SS \dots S}_{x \text{ times}} \underbrace{FF \dots F}_{n-x \text{ times}}\}$$

and all its possible **combinations**, each with probability $p^x(1-p)^{n-x}$

Remark (Reminder of Combinations)

Recall (see Intro lecture) that combinations are defined as:

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} = C_n^k$$

and, for $n \geq k$, we say "n choose k".

The binomial coefficient $\binom{n}{k}$ represents the number of possible combinations of n objects taken k at a time, *when the* order doesn't matter*.*

Thus, C_n^k represents the number of different groups of size k that could be selected from a set of n objects when the order of selection is not relevant.

Binomial

Definition (Binomial PMF)

$$
\begin{aligned}
P(\{X = x\}) &= \binom{n}{x} p^x (1-p)^{n-x} \qquad (1) \\
&= \frac{n!}{x!\,(n-x)!} p^x (1-p)^{n-x}, \text{ for } x = 0, 1, 2, ..., n \qquad (2)
\end{aligned}
$$

Remark (Interpretation of the Formula)

1. *The first factor*
$$\binom{n}{k} = \frac{n!}{x!\,(n-x)!}$$
is the number of different combinations of individual 'successes' and 'failures' in n (Bernoulli) trials that result in a sequence containing a total of x 'successes' and $n-x$ 'failures'.

2. *The second factor*
$$p^x (1-p)^{n-x}$$
is the probability associated with any one sequence of x 'successes' and $(n-x)$ 'failures'.

Binomial

- Mean:

$$\begin{aligned} E[X] &= \sum_{x=0}^{n} xP\{X = x\} \\ &= \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} = np \end{aligned}$$

Binomial

- Mean:

$$
\begin{aligned}
E[X] &= \sum_{x=0}^{n} xP\{X = x\} \\
&= \sum_{x=0}^{n} x \binom{n}{x} p^{x} (1-p)^{n-x} = np
\end{aligned}
$$

- Variance:

$$
\begin{aligned}
Var(X) &= \sum_{x=0}^{n} (x - np)^{2} P(\{X = x\}) \\
&= np(1-p)
\end{aligned}
$$

Binomial

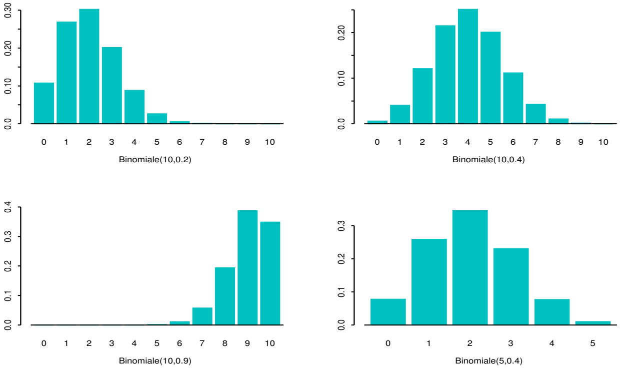
Remark (Short-hand notation)

$$X \sim \mathcal{B}(n, p)$$

Remark

- *The Bernoulli distribution is a special case, when* $n = 1$ *i.e.* $\mathcal{B}(1, p)$
- *Roughly speaking, "a Binomial random variable arises when we sum* n *independent Bernoulli trails."*

Binomial


0.0
0.10
0.20
0.30
0 1 2 3 4 5 6 7 8 9 10
Binomiale(10,0.2)
0.0
0.10
0.20
0 1 2 3 4 5 6 7 8 9 10
Binomiale(10,0.4)
0.0
0.1
0.2
0.3
0.4
0 1 2 3 4 5 6 7 8 9 10
Binomiale(10,0.9)
0.0
0.1
0.2
0.3
0 1 2 3 4 5
Binomiale(5,0.4)

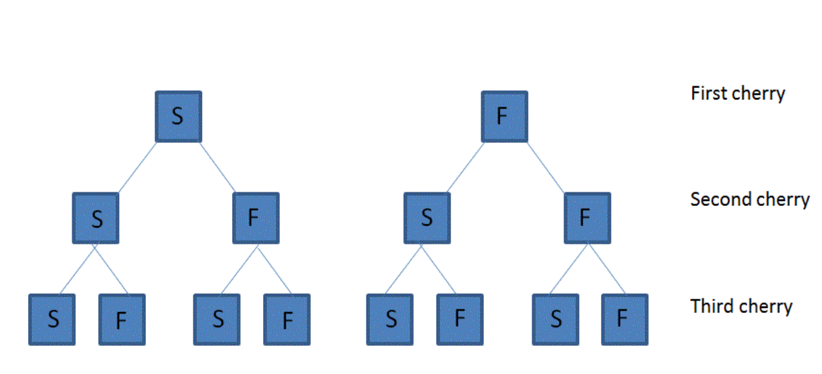
Example (cherry trees)

One night a storm washes three cherries ashore on an island. For each cherry, there is a probability $p = 0.8$ that its seed will produce a tree. What is the probability that these three cherries will produce two trees?

First, we notice that each of these events can be characterised using a **Bernoulli distribution**. To this end, consider whether each seed will produce a tree as a sequence of $n = 3$ trials. For each cherry:

- either the cherry produces a tree (Success) or it does not (Failure);
- the event that a cherry produces a tree is independent from the event that any of the other two cherries produces a tree.
- The probability that a cherry produces a tree is the same for all three cherries

Example (continued)



Binomial

Example (continued)

- There are $2^3 = 8$ possible outcomes from the 3 individual trials

Binomial

Example (continued)

- There are $2^3 = 8$ possible outcomes from the 3 individual trials
- It does not matter which of the three cherries produce a tree

Example (continued)

- There are $2^3 = 8$ possible outcomes from the 3 individual trials
- It does not matter which of the three cherries produce a tree
- Consider all of the possible sequences of outcomes (S=success, F=failure)

SSS, SSF, SFS, SFF, FSS , FSF, FFS, FFF

Binomial

Example (continued)

- There are $2^3 = 8$ possible outcomes from the 3 individual trials
- It does not matter which of the three cherries produce a tree
- Consider all of the possible sequences of outcomes (S=success, F=failure)

SSS, SSF, SFS, SFF, FSS , FSF, FFS, FFF

- We are interested in SSF, SFS, FSS

Example (continued)

- There are $2^3 = 8$ possible outcomes from the 3 individual trials
- It does not matter which of the three cherries produce a tree
- Consider all of the possible sequences of outcomes (S=success, F=failure)

SSS, SSF, SFS, SFF, FSS , FSF, FFS, FFF

- We are interested in SSF, SFS, FSS
- These possible events are *mutually exclusive*, so

$$P(\{SSF \cup SFS \cup FSS\}) = P(\{SSF\}) + P(\{SFS\}) + P(\{FSS\})$$

Example (continued)

The three trials are assumed to be *independent*, so each of the three seed events corresponding to two trees growing has the same probability

$$
\begin{aligned}
P(\{SSF\}) &= P(\{S\}) \cdot P(\{S\}) \cdot (P\{F\}) \\
&= 0.8 \cdot 0.8 \cdot (1-0.8) \\
&= 0.8 \cdot (1-0.8) \cdot 0.8 = P(\{SFS\}) \\
&= (1-0.8) \cdot 0.8 \cdot 0.8 = P(\{FSS\}) \\
&= 0.128
\end{aligned}
$$

So the probability of two trees resulting from the three seeds must be

$$
\begin{aligned}
P(\{SSF \cup SFS \cup FSS\}) &= 3 \cdot 0.128 \\
&= 0.384.
\end{aligned}
$$

Binomial

Example

Finally, we notice that we can obtain the same result (in a more direct way), using the **binomial probability** for the random variable

$$X = \text{number of trees that grow from 3 seeds.}$$

Indeed

$$\begin{aligned} P(\{X=2\}) &= \frac{3!}{2!\,(3-2)!} \cdot (0.8)^2 \cdot (1-0.8)^{3-2} \\ &= 3 \cdot (0.8)^2 \cdot (0.2) \\ &= 0.384. \end{aligned}$$

Poisson

Poisson

Definition

Let us consider random variable X which takes values $0, 1, 2, \ldots$, i.e. elements of $\mathbb{N}$.
X is said to be a Poisson random variable if its probability mass function, with $\lambda > 0$ fixed and providing info on the *intensity*, is:

$$p(x) = P(\{X = x\}) = \frac{\lambda^x e^{-\lambda}}{x!}, \qquad x = 0, 1, 2, \ldots \tag{3}$$

and we write $X \sim \mathcal{P}(\lambda)$.

The Eq. (3) defines a genuine probability mass function, since $p(x) \geq 0$ and

$$\begin{aligned}\sum_{x=0}^{\infty} p(x) &= \sum_{x=0}^{\infty} \frac{\lambda^x e^{-\lambda}}{x!} \\ &= e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \\ &= e^{-\lambda} e^{\lambda} = 1 \quad \text{(see Intro Lecture).}\end{aligned}$$

Poisson

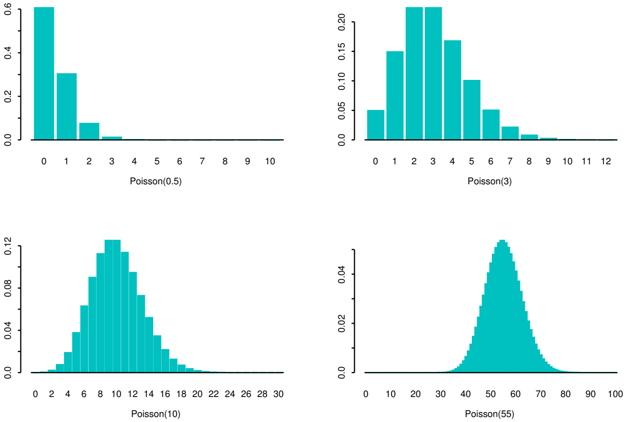
Moreover, for a given value of λ also the CDF can be easily defined. E.g.

$$F_X(2) = P(\{X \leq 2\}) = e^{-\lambda} + \lambda e^{-\lambda} + \frac{\lambda^2 e^{-\lambda}}{2},$$

and the Expected value and Variance for Poisson distribution (see tutorial) can be obtained by "sum algebra" (and/or some algebra)

$$\begin{aligned} E[X] &= \lambda \\ Var(X) &= \lambda. \end{aligned}$$

Poisson


0.0
0.2
0.4
0.6
0 1 2 3 4 5 6 7 8 9 10
Poisson(0.5)
0.0
0.05
0.10
0.15
0.20
0 1 2 3 4 5 6 7 8 9 10 11 12
Poisson(3)
0.0
0.04
0.08
0.12
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30
Poisson(10)
0.0
0.02
0.04
0 10 20 30 40 50 60 70 80 90 100
Poisson(55)

Example

The average number of newspapers sold by Alfred is 5 per minute[a]*. What is the probability that Alfred will sell at least 1 newspaper in a minute?*

Poisson

Example

The average number of newspapers sold by Alfred is 5 per minute[a]*. What is the probability that Alfred will sell at least 1 newspaper in a minute?*

To answer, let X be the # of newspapers sold by Alfred in a minute. We have

$$X \sim \mathcal{P}(\lambda)$$

with $\lambda = 5$, so

$$\begin{aligned} P(X \geq 1) &= 1 - P(\{X = 0\}) \\ &= 1 - \exp^{-5} \frac{5^0}{0!} \\ &\approx 1 - 0.0067 \approx 99.33\%. \end{aligned}$$

How about $P(X \geq 2)$? Is it $P(X \geq 2) \geq P(X \geq 1)$ or not? Answer the question...

[a] This number provides info on the intensity at which a random phenomenon occurs.

Example

A telephone switchboard handles 300 calls, on the average, during one hour. The board can make maximum 10 connections per minute. Use the Poisson distribution to evaluate the probability that the board will be overtaxed during a given minute.

To answer, let us set $\lambda = 300$ per hour, which is equivalent to 5 calls per minute. Noe let us define

$$X = \# \text{ of connections in a minute}$$

and by assumption we have $X \sim \mathcal{P}(\lambda)$. Thus,

$$\begin{aligned} P[\text{overtaxed}] &= P(\{X > 10\}) \\ &= 1 - \underbrace{P(\{X \leq 10\})}_{\text{using } \lambda = 5, \text{ minute base}} \\ &\approx 0.0137. \end{aligned}$$

Remark

Let us consider $X \sim \mathcal{B}(n, p)$, where n is large, p is small, and the product np is noticeable. Setting, $\lambda = np$, we then have that, for the Binomial probability as in Eq.(2), it is a good approximation to write:

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

Remark

Let us consider $X \sim \mathcal{B}(n, p)$, where n is large, p is small, and the product np is noticeable. Setting, $\lambda = np$, we then have that, for the Binomial probability as in Eq.(2), it is a good approximation to write:

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

Proof.

To see this, remember that

$$\lim_{n \to \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}.$$

Then, let us consider that in our setting, we have $p = \lambda/n$. From the formula of the binomial probability mass function we have:

$$p(0) = (1 - p)^n = \left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda}, \quad \text{as} \quad n \to \infty.$$

$\square$

Proof.

Moreover, it is easily found that

$$\frac{p(k)}{p(k-1)} \quad = \quad \frac{np-(k-1)p}{k(1-p)} \approx \frac{\lambda}{k}, \quad \text{as} \quad n \to \infty.$$

Therefore, we have

$$\begin{array}{rcl} p(1) & \approx & \frac{\lambda}{1!}p(0) \approx \lambda e^{-\lambda} \\ p(2) & \approx & \frac{\lambda}{2!}p(1) \approx \frac{\lambda^2}{2}e^{-\lambda} \\ \cdots & \cdots & \cdots \\ p(k) & \approx & \frac{\lambda}{k!}p(k-1) \approx \underbrace{\frac{\lambda^k}{k!}e^{-\lambda}}_{\text{see} \quad \text{Eq. (3)}} \end{array}$$

thus, we remark that $p(k)$ can be approximated by the probability mass function of a Poisson — which is easier to evaluate. $\square$

Poisson

Example (two-fold use of Poisson)

Suppose a certain high-speed printer makes errors at random on printed paper[a]*.*
Assuming that the Poisson distribution with parameter $\lambda = 4$ is appropriate to model the number of errors per page (say, X), what is the probability that in a book containing 300 pages (produced by the printer) at least 7 will have no errors?

Let X denote the number of errors per page, so that

$$p(x) = \exp(-4)\frac{4^x}{x!}, \quad \text{for} \quad x = 0, 1, 2, \ldots.$$

The probability of any page to be error free is then

$$p(0) = \exp(-4)\frac{4^0}{0!} = \exp(-4) \approx 0.018.$$

[a] This exercise is related to Ex 2 of PS6. The calculation is similar but not identical: notice the difference between the size of p in this example and in the tutorial.

Example (cont'd)

Having no errors on a page is a success, and there are 300 independent pages. Hence, let us define

$$Y = \text{the number of pages without any errors.}$$

Y is binomially distributed with parameters $n = 300$ and $p = 0.018$, namely

$$Y \sim \mathcal{B}(n, p).$$

But here we have

$$n \text{ large, } p \text{ small, and } np = 5.4$$

thus, we can compute $P(\{Y \geq 7\})$ using either the exact Binomial or its Poisson approximation. So

- using $\mathcal{B}(300, 0.018)$, we get: $P(\{Y \geq 7\}) \approx 0.297$
- using $\mathcal{P}(5.4)$, we get $P(\{Y \geq 7\}) \approx 0.298$.

Hypergeometric

Definition

Let us consider a random experiment consisting of a series of n trials, having the following properties

1. Only two mutually exclusive outcomes are possible in each trial: success (S) and failure (F)

The random variable

$$X = \text{number of successes in } n \text{ such trials}$$

has an hypergeometric distribution and ...

Hypergeometric

Definition

Let us consider a random experiment consisting of a series of n trials, having the following properties

1. Only two mutually exclusive outcomes are possible in each trial: success (S) and failure (F)
2. The population has N elements with k Successes S and $N - k$ Failures F

The random variable

$$X = \text{number of successes in } n \text{ such trials}$$

has an hypergeometric distribution and ...

Definition

Let us consider a random experiment consisting of a series of n trials, having the following properties

1. Only two mutually exclusive outcomes are possible in each trial: success (S) and failure (F)
2. The population has N elements with k Successes S and $N - k$ Failures F
3. Sampling from the population is done **without** replacement (so that the **trials are not independent**).

The random variable

$$X = \text{number of successes in } n \text{ such trials}$$

has an hypergeometric distribution and ...

Definition (cont'd)

... the probability that $X = x$ is

$$P(\{X = x\}) = \frac{\begin{pmatrix} k \\ x \end{pmatrix} \begin{pmatrix} N-k \\ n-x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}.$$

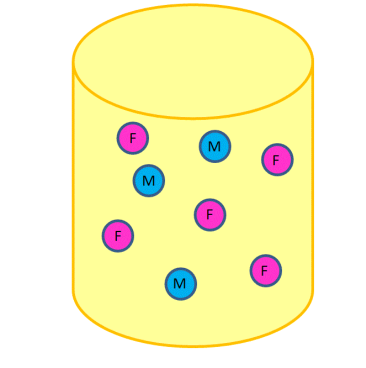
Moreover,

$$\begin{aligned} E\left[X\right] &= \frac{nk}{N} \\ Var\left(X\right) &= \frac{nk\left(N-k\right)\left(N-n\right)}{N^2\left(N-1\right)} \end{aligned}$$

Hypergeometric

Example (Psychological experiment)

A group of 8 students includes 5 women and 3 men: 3 students are randomly chosen to participate in a psychological experiment. What is the probability that *exactly* 2 women will be included in the sample?



Example (cont'd)

Consider each of the three participants being selected as a separate trial $\Rightarrow$ there are $n = 3$ trials. Consider a woman being selected in a trial as a 'success'
Then here $N = 8$, $k = 5$, $n = 3$, and $x = 2$, so that

$$
\begin{aligned}
P(\{X = 2\}) &= \frac{\begin{pmatrix} 5 \\ 2 \end{pmatrix} \begin{pmatrix} 8-5 \\ 3-2 \end{pmatrix}}{\begin{pmatrix} 8 \\ 3 \end{pmatrix}} \\
&= \frac{\frac{5!}{2!3!} \frac{3!}{1!2!}}{\frac{8!}{5!3!}} \\
&= 0.53571
\end{aligned}
$$

Negative Binomial

Negative Binomial

Definition

Let us consider a random experiment consisting of a series of trials, having the following properties

1. Only two mutually exclusive outcomes are possible in each trial: 'success' (S) and 'failure' (F)

What is the probability of having exactly y F's before the r^{th} S?

Equivalently: What is the probability that in a sequence of $y + r$ (Bernoulli) trials the last trial yields the r^{th} S?

Negative Binomial

Definition

Let us consider a random experiment consisting of a series of trials, having the following properties

1. Only two mutually exclusive outcomes are possible in each trial: 'success' (S) and 'failure' (F)
2. The outcomes in the series of trials constitute *independent events*

What is the probability of having exactly y F's before the r^{th} S?

Equivalently: What is the probability that in a sequence of $y + r$ (Bernoulli) trials the last trial yields the r^{th} S?

Negative Binomial

Definition

Let us consider a random experiment consisting of a series of trials, having the following properties

1. Only two mutually exclusive outcomes are possible in each trial: ‘success’ (S) and ‘failure’ (F)
2. The outcomes in the series of trials constitute *independent events*
3. The probability of success p in each trial is *constant* from trial to trial

What is the probability of having exactly y F’s before the r^{th} S?

Equivalently: What is the probability that in a sequence of $y + r$ (Bernoulli) trials the last trial yields the r^{th} S?

Definition

Let

$X =$ the total number of trials required until a total of r successes is accumulated.

Then X is said to be a Negative Binomial random variable and its probability mass function $P(\{X = n\})$ equals the probability of $r-1$ 'successes' in the first $n-1$ trials, times the probability of a 'success' on the last trial. These probabilities are given by

$$P(\{X = n\}) = \left(\begin{array}{c} n-1 \\ r-1 \end{array} \right) p^r (1-p)^{n-r} \qquad \text{for } n = r, r+1, \ldots$$

The mean and variance for X are, respectively,

$$\begin{aligned} E[X] &= \frac{r}{p} \\ Var(X) &= \frac{r(1-p)}{p^2} \end{aligned}$$

Example (marketing research)

- A marketing researcher wants to find 5 people to join her focus group

Example (marketing research)

- A marketing researcher wants to find 5 people to join her focus group
- Let p denote the probability that a randomly selected individual agrees to participate in the focus group

Example (marketing research)

- A marketing researcher wants to find 5 people to join her focus group
- Let p denote the probability that a randomly selected individual agrees to participate in the focus group
- If $p = 0.2$, what is the probability that the researcher must ask 15 individuals before 5 are found who agree to participate?

Negative Binomial

Example

Example (marketing research)

- A marketing researcher wants to find 5 people to join her focus group
- Let p denote the probability that a randomly selected individual agrees to participate in the focus group
- If $p = 0.2$, what is the probability that the researcher must ask 15 individuals before 5 are found who agree to participate?
- In this case, $p = 0.2$, $r = 5$, $n = 15$: we are looking for $P(\{X = 15\})$. By the negative binomial formula we have

$$
\begin{aligned}
P(\{X = 15\}) &= \binom{14}{4} (0.2)^5 (0.8)^{10} \\
&= 0.034
\end{aligned}
$$

Geometric

Definition (a special case)

When $r = 1$, the negative binomial distribution is equivalent to the **Geometric distribution** — see the example in the previous lecture.

In this case, probabilities are given by

$$P(\{X = n\}) = p\,(1-p)^{n-1}\,,\ \text{for } n = 1, 2, \ldots$$

The corresponding mean and variance for X are, respectively,

$$\begin{aligned} E[X] &= \frac{1}{p} \\ Var(X) &= \frac{(1-p)}{p^2} \end{aligned}$$

Example (failure of a machine)

Items are produced by a machine having a 3% defective rate.

- What is the probability that the first defective occurs in the fifth item inspected?

$$\begin{aligned} P(\{X=5\}) &= P(\text{first 4 non-defective})P(\text{5th defective}) \\ &= (0.97)^4(0.03) \approx 0.026 \end{aligned}$$

- What is the probability that the first defective occurs in the first five inspections?

$$\begin{aligned} P(\{X \leq 5\}) = P(\{X < 6\}) &= P(\{X=1\}) + \ldots + P(\{X=5\}) \\ &= 1 - P(\text{first 5 non-defective}) = 0.1412. \end{aligned}$$

More generally, for a geometric random variable we have:

$$P(\{X \geq k\}) = (1-p)^{k-1}.$$

Thus, in the example we have $P(\{X \geq 6\}) = (1-0.03)^{6-1} \approx 0.8587$

$$P(\{X \leq 5\}) = 1 - P(\{X \geq 6\}) \approx 1 - 0.8587 \approx 0.1412.$$

Wrap-up

It is important to keep in mind and review:

- The expressions of the PMF's of the different functions.
- Their case uses (When to use Bernoulli, Binomial, Poisson, ...)
- The link between Poisson and Binomial, as it simplifies the computations.

Thank You for your Attention!

Thank You for your Attention!

"See you" Next Week