Mathematical Statistics

MAS 713

Chapter 3.2

Previous lectures

1. Random variables
2. Discrete Random Variables
3. Continuous Random Variables
4. Expectation of a random variable
5. Variance of a random variable

Questions?

This lecture

3.2. Special random variables

- 3.2.1 Introduction
- 3.2.2 Binomial random variables
- 3.2.3 Hypergeometric random variables
- 3.2.4 Poisson random variables
- 3.2.5 Uniform random variables
- 3.2.6 Exponential random variables

Additional reading : Chapter 3 in the textbook

Introduction

In practice, we are often running same kind of (random) experiments, from which we are often interested in the same kind of results

$\rightsquigarrow$ it turns out that certain “types” of random variables come up over and over again in applications

In this chapter, we’ll study a variety of those **special random variables**

You can also go to

http://www.socr.ucla.edu/htmls/dist/

and have a look at the numerous ‘special’ distributions there

Exponential Family

Exponential Family - Background

Definition

- A family of discrete probabilities $(\mathbb{P}_\theta)_\theta$ is an exponential family if each of its probability mass function $p(x;\boldsymbol{\theta})$ can be represented in the following General Form:

$$p(x;\boldsymbol{\theta}) = h(\boldsymbol{x})g(\boldsymbol{\theta})\exp\left(\boldsymbol{\theta}^T u(\boldsymbol{x})\right)$$

- A family of continuous probabilities $(\mathbb{P}_\theta)_\theta$ is an exponential family if each of its probability density function $f(x;\boldsymbol{\theta})$ can be represented in the following General Form:

$$f(x;\boldsymbol{\theta}) = h(\boldsymbol{x})g(\boldsymbol{\theta})\exp\left(\boldsymbol{\theta}^T u(\boldsymbol{x})\right)$$

- $\boldsymbol{\theta}$ - vector of the natural parameters
- $u(\boldsymbol{x}), h(\boldsymbol{x})$ - some function of $\boldsymbol{x}$.
- $g(\boldsymbol{\theta})$ - normalizes the distribution $g(\boldsymbol{\theta})\int h(\boldsymbol{x})\exp\left(\boldsymbol{\theta}^T \boldsymbol{u}(\boldsymbol{x})\right) d\boldsymbol{x} = 1$

- **Example:** Bernoulli Distribution -

$$\begin{aligned}
Bernoulli(x|\pi) := p_{Ber}(x;\pi) &= \pi^x(1-\pi)^{1-x} \\
&= \exp\left(x\ln\pi + (1-x)\ln(1-\pi)\right) \\
&= \exp\left(x\left(\ln\pi - \ln(1-\pi)\right) + \ln(1-\pi)\right) \\
&= \exp\left(x\ln\left(\frac{\pi}{1-\pi}\right)\right)\exp\left(\ln(1-\pi)\right) \\
&= (1-\pi)\exp\left(\ln\left(\frac{\pi}{1-\pi}\right)x\right)
\end{aligned}$$

- Setting $\theta := \ln\left(\frac{\pi}{1-\pi}\right)$ which gives the *logistic sigmoid* function $\pi = \sigma(\theta) := \frac{1}{1+\exp(-\theta)}$
- $u(x) = x$; $h(x) = 1$; $g(\theta) = \sigma(-\pi)$.

$$\Longrightarrow p_{Ber}(x;\theta) = h(x)g(\theta)\exp\left(\theta^T u(x)\right)$$

EXAMPLES OF DISCRETE MEMBERS:

Bernoulli:	$p(x;\pi) = \pi^x (1-\pi)^{1-x}$	$x \in \{0,1\}$
Binomial:	$p(x;n,\pi) = C_r^n \pi^x (1-\pi)^{n-x}$	$x \in \{0,1,\dots,n\}$
Multinomial:	$p(x;\boldsymbol{\pi}) = \frac{n!}{x_1!x_2!\dots x_n!} \prod_{i=1}^n \pi_i^{x_i}$	$x_i \in \{0,1,\dots,n\}$
Poisson:	$p(x;\lambda) = \frac{\lambda^x}{x!} e^{-\lambda}$	$x \in \mathbb{N}_0$
Dirichlet:	$p(x;\boldsymbol{\pi}) = \frac{\Gamma\left(\sum_i \pi_i\right)}{\prod_i \Gamma(\pi_i)} \prod_i x_i^{\pi_i - 1}$	$x \in [0,1], \sum_i x_i = 1$

EXAMPLES OF DISCRETE MEMBERS:

Binomial Model

- used to model the number of successes in a sequence of n independent Bernoulli yes/no events.

EXAMPLES OF DISCRETE MEMBERS:

Multinomial Model

- analog of the Bernoulli distribution for the categorical distribution
- each trial results in exactly one of some fixed finite number k of possible outcomes, with probabilities $p_1, \ldots, p_k$.
- There are still n - independent trials.

EXAMPLES OF DISCRETE MEMBERS:

Poisson Model

- captures the probability of the number of events occurring in a fixed period of time.

- It assumes these events occur with a known average rate and independently of the time since the last event.

- It can also be used for the number of events in intervals such as distance, area or volume.

EXAMPLES OF CONTINUOUS MEMBERS:

Gaussian:	$p(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	$x \in \mathbb{R}$
Inverse Normal:	$p(x;\lambda,\mu) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left(-\frac{\lambda}{2\mu^2 x}(x-\mu)^2\right)$	$x \in \mathbb{R}^+$
Exponential:	$p(x;\lambda) = \lambda \exp(-\lambda x)$	$x \in \mathbb{R}^+$
Laplace:	$p(x;a,b) = \frac{1}{2b} \exp(-\frac{\lvert x-\mu\rvert}{b})$	$x \in \mathbb{R}$
Beta:	$p(x;\pi,\beta) = \frac{y^{\pi-1}(1-x)^{\beta-1}}{B(\pi,\beta)}$	$x \in [0,1]$
Gamma:	$p(x;k,\theta) = x^{k-1}\frac{\exp\left(-\frac{x}{\theta}\right)}{\Gamma(k)\theta^k}$	$x \in \mathbb{R}^+$

EXAMPLES OF CONTINUOUS MEMBERS:

Exponential Model

- describes the inter arrival time of events in a homogeneous Poisson process.

Used in practical models to approximate:

- time it takes until a radio active particle decays.
- time taken until received next phone call at exchange.
- time until a default occurs in credit modeling, or loss is recorded in Operational Risk.

EXAMPLES OF CONTINUOUS MEMBERS:

Beta Model

- often used as a prior distribution on a random variable on the interval [0,1] - ie. a probability.

EXAMPLES OF CONTINUOUS MEMBERS:

Gamma Model

- often used as a distribution of waiting times or failure times.

Independent random variables

Definition

Random variables X and Y are said to be independent **if and only if** for all $(x, y) \in \mathbb{R} \times \mathbb{R}$,

$$\mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x) \times \mathbb{P}(Y \leq y)$$

- X, Y indep. $\iff$ all couples of events $(X \leq x)$, $(Y \leq y)$ indep.

Property

Random variables X and Y are said to be independent **if and only if** for all functions h, g

$$\mathbb{E}(h(X)g(Y)) = \mathbb{E}(h(X)) \times \mathbb{E}(g(Y))$$

Example: X and Y independent $\implies$ $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$.

Examples

Consider the following random experiments and random variables :

- flip a coin 10 times. Let $X =$ number of heads obtained.
- a worn machine tool produces 1% defective parts.
  Let $X =$ number of defective parts in the next 25 parts produced.
- each sample of air has a 10% chance of containing a particular molecule.
  Let $X =$ the number of air samples that contain the molecule in the next 18 samples analysed.
- a multiple-choice test contains 10 questions, each with 4 choices, and you guess at each question.
  Let $X =$ the number of questions answered correctly.

$\rightsquigarrow$ similar experiments, **similar** random variables

$\rightsquigarrow$ a general framework would be very useful

The Binomial distribution

The Binomial distribution

Assume:

- the outcome of a random experiment can be classified as either a "Success" (S) or a "Failure" (F) ($\rightsquigarrow S = \{\text{Success}, \text{Failure}\}$)
- we observe a Success with probability π
- n independent repetitions of this experiment are performed

Define X = number of successes observed over the n repetitions.
We say that X is a binomial random variable with parameters n and π :

$$\boxed{X \sim \text{Bin}(n, \pi)}$$

See that $S_X = \{0, 1, 2, \ldots, n\}$ and the binomial probability mass is

$$p_{Bin}(x) = \binom{n}{x} \pi^x (1-\pi)^{n-x} \qquad \text{for } x \in S_X$$

where $\binom{n}{x}$ is the number of different groups of x objects that can be chosen from a set of n objects

The Binomial distribution

Note : the coefficients $\binom{n}{x} = n!/(x!(n-x)!)$ are called the binomial coefficients, they are the coefficients arising in the famous Newton's binomial expansion

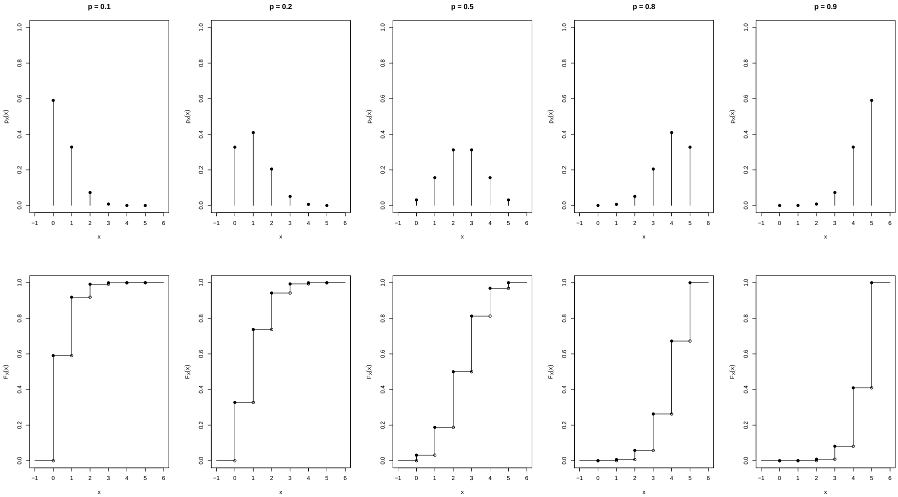
$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

These coefficients are often represented in the Pascal's triangle, named after the French mathematician Blaise Pascal (1623-1662)

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1
1 7 21 35 35 21 7 1
1 8 28 56 70 56 28 8 1
1 9 36 84 126 126 84 36 9 1
1 10 45 120 210 252 210 120 45 10 1
1 11 55 165 330 462 462 330 165 55 11 1
1 12 66 220 495 792 924 792 495 220 66 12 1
1 13 78 186 715 1287 1716 1716 1287 715 186 78 13 1

The Binomial distribution : pmf and cdf

Binomial pmf and cdf, for $n = 5$ and $\pi = \{0.1, 0.2, 0.5, 0.8, 0.9\}$



The Bernoulli distribution

The Bernoulli distribution

Special case $\text{Bin}(1,\pi) \leadsto$ the Bernoulli distribution,

$$\boxed{X \sim \text{Ber}(\pi)}$$

pmf :

$$p_{Ber}(x) = \begin{cases} 1-\pi & \text{if } x = 0 \\ \pi & \text{if } x = 1 \\ 0 & \text{otherwise} \end{cases}$$

Note : if $X \sim \text{Bin}(n,\pi)$, we can represent it as

$$X = \sum_{i=1}^{n} X_i$$

where X_i's are n **independent** Bernoulli r.v. with parameters π

$\leadsto$ each repetition of the experiment in the Binomial framework is called a Bernoulli trial

Binomial distribution : properties

- First note that

$$\sum_{x \in S_X} p(x) = \sum_{x=0}^{n} \binom{n}{x} \pi^x (1-\pi)^{n-x} = (\pi + (1-\pi))^n = 1$$

using the binomial expansion

- Second, it is easy to see that
if $X_1 \sim \text{Bin}(n_1, \pi)$, $X_2 \sim \text{Bin}(n_2, \pi)$ and X_1 is independent of X_2, then

$$X_1 + X_2 \sim \text{Bin}(n_1 + n_2, \pi)$$

Binomial distribution : expectation and variance

- Remind the representation $X = \sum_{i=1}^{n} X_i$, with $X_i \sim \text{Bern}(\pi)$
- We know that

$$\mathbb{E}(X_i) = \pi \qquad \text{and} \qquad \mathbb{V}\text{ar}(X_i) = \pi(1-\pi)$$

- It follows

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{E}(X_i) = \sum_{i=1}^{n} \pi = n\pi,$$

-

$$\mathbb{V}\text{ar}(X) = \mathbb{V}\text{ar}\left(\sum_{i=1}^{n} X_i\right) \overset{\text{ind.}}{=} \sum_{i=1}^{n} \mathbb{V}\text{ar}(X_i) = \sum_{i=1}^{n} \pi(1-\pi) = n\pi(1-\pi)$$

Binomial distribution : expectation and variance

Mean and variance of the binomial distribution

If $X \sim \text{Bin}(n, \pi)$,

$$\mu = \mathbb{E}(X) = n\pi \qquad \text{and} \qquad \sigma^2 = \mathbb{V}\text{ar}(X) = n\pi(1-\pi)$$

Binomial distribution : examples

Example

It is known that disks produced by a certain company will be defective with probability 0.01 independently of each other. The company sells the disk in packages of 10 and offers a money-back guarantee that at most 1 of the disks is defective.

a) In the long-run, what proportion of packages is returned?
b) If someone buys three packages, what is the probability that exactly one of them will be returned?

a) Let X be the number of defective disks in a package. Then, it is clear that

$$X \sim \text{Bin}(10, 0.01)$$

Hence
$$\begin{aligned}\mathbb{P}(X > 1) &= 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) \\ &= 1 - \binom{10}{0} 0.01^0 0.99^{10} - \binom{10}{1} 0.01^1 0.99^9 \simeq 0.005\end{aligned}$$

$\rightsquigarrow$ in the long-run, 0.5 percent of the packages will have to be replaced

Binomial distribution : examples

Example

It is known that disks produced by a certain company will be defective with probability 0.01 independently of each other. The company sells the disk in packages of 10 and offers a money-back guarantee that at most 1 of the disks is defective.

a) In the long-run, what proportion of packages is returned?
b) If someone buys three packages, what is the probability that exactly one of them will be returned?

a) Let X be the number of defective disks in a package. Then, it is clear that

$$X \sim \text{Bin}(10, 0.01)$$

Hence
$$\begin{aligned}\mathbb{P}(X > 1) &= 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) \\ &= 1 - \binom{10}{0} 0.01^0 0.99^{10} - \binom{10}{1} 0.01^1 0.99^9 \simeq 0.005\end{aligned}$$

$\rightsquigarrow$ in the long-run, 0.5 percent of the packages will have to be replaced

Example

It is known that disks produced by a certain company will be defective with probability 0.01 independently of each other. The company sells the disk in packages of 10 and offers a money-back guarantee that at most 1 of the disks is defective.

a) What proportion of packages is returned?
b) If someone buys three packages, what is the probability that exactly one of them will be returned?

b) Let Y be the number of packages that the person has to return. We have

$$Y \sim \text{Bin}(3, \pi)$$

where π is the probability that a package is returned, that is, contains more than 1 defective disk.

- From a), we know that $\pi = 0.005$

Thus, the probability that exactly one of the three packages is returned equals

$$\mathbb{P}(Y = 1) = \binom{3}{1} 0.005^1 0.995^2 = 0.015$$

Example

It is known that disks produced by a certain company will be defective with probability 0.01 independently of each other. The company sells the disk in packages of 10 and offers a money-back guarantee that at most 1 of the disks is defective.

a) What proportion of packages is returned?
b) If someone buys three packages, what is the probability that exactly one of them will be returned?

b) Let Y be the number of packages that the person has to return. We have

$$Y \sim \text{Bin}(3, \pi)$$

where π is the probability that a package is returned, that is, contains more than 1 defective disk.

- From a), we know that $\pi = 0.005$

Thus, the probability that exactly one of the three packages is returned equals

$$\mathbb{P}(Y = 1) = \binom{3}{1} 0.005^1 0.995^2 = 0.015$$

Example

Suppose that 10% of all bits transmitted through a digital communication channel are erroneously received and that whether any is erroneously received is independent of whether any other bit is erroneously received. Consider sending a large number of messages, each consisting of 20 bits.

a) What proportion of these messages will have exactly 2 erroneously bits?
b) What proportion of these messages will have at least 5 erroneously bits?
c) For what proportion of these messages will more than half the bits be erroneously?

Let X be the number of erroneously received bits in a message of 20 bits.

- Clearly, we have $X \sim \text{Bin}(20, 0.1)$. Thus we have

a) $\mathbb{P}(X = 2) = \binom{20}{2} 0.1^2 0.9^{18} = 0.2852$

b) $\mathbb{P}(X \geq 5) = 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) - \mathbb{P}(X = 2) - \mathbb{P}(X = 3) - \mathbb{P}(X - 4) = \ldots$

c) $\mathbb{P}(X > 10) = \mathbb{P}(X = 11) + \mathbb{P}(X = 12) + \ldots + \mathbb{P}(X = 20) = \ldots$

$\rightsquigarrow$ very tedious ! $\rightsquigarrow$ use **statistical tables** or a statistical software

Example

Suppose that 10% of all bits transmitted through a digital communication channel are erroneously received and that whether any is erroneously received is independent of whether any other bit is erroneously received. Consider sending a large number of messages, each consisting of 20 bits.

a) What proportion of these messages will have exactly 2 erroneously bits?
b) What proportion of these messages will have at least 5 erroneously bits?
c) For what proportion of these messages will more than half the bits be erroneously?

Let X be the number of erroneously received bits in a message of 20 bits.

- Clearly, we have $X \sim \text{Bin}(20, 0.1)$. Thus we have

a) $\mathbb{P}(X = 2) = \binom{20}{2} 0.1^2 0.9^{18} = 0.2852$

b) $\mathbb{P}(X \geq 5) = 1 - \mathbb{P}(X = 0) - \mathbb{P}(X = 1) - \mathbb{P}(X = 2) - \mathbb{P}(X = 3) - \mathbb{P}(X - 4) = \ldots$

c) $\mathbb{P}(X > 10) = \mathbb{P}(X = 11) + \mathbb{P}(X = 12) + \ldots + \mathbb{P}(X = 20) = \ldots$

$\rightsquigarrow$ very tedious ! $\rightsquigarrow$ use **statistical tables** or a statistical software

Binomial distribution : examples

Command Window

New to MATLAB? Watch this Video, see Demos, or read Getting Started.

```
>> binopdf(2,20,0.1)

ans =

    0.2852

>> 1-binocdf(4,20,0.1)

ans =

    0.0432

>> 1-binocdf(10,20,0.1)

ans =

  7.0886e-007

>>
```

The Hypergeometric distribution

The Hypergeometric distribution

Consider the following situation :

We are interested in the number of defectives in a sample of n units drawn from a lot containing N units, of which k are defective

Careful: It can be tempting to regard X, the number of defectives in the sample, as a Binomial random variable : n units are repeatedly classified as defective ('Success') or not defective ('Failure')

However, if the first drawing yields a defective with probability k/N, the second drawing does with probability $(k-1)/(N-1)$ or $k/(N-1)$ (depending on the first drawing)

$\leadsto$ the 'trials' are not independent, and the probability of success is not constant $\leadsto$ violation of the Binomial distribution assumptions

$\leadsto$ X actually follows the so-called Hypergeometric distribution

$$\boxed{X \sim \text{Hyp}(N, k, n)}$$

Hypergeometric distribution : properties

- See that $S_X = \{\max(0, n+k-N), \ldots, \min(k, n)\}$ and the hypergeometric probability mass function is given by

$$p(x) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} \qquad \text{for } x \in S_X$$

- The mean and variance of the Hypergeometric distribution can be determined from the trials that comprise the experiment
- **However**, the trials are not independent, and so the calculations a more difficult than for the Binomial.

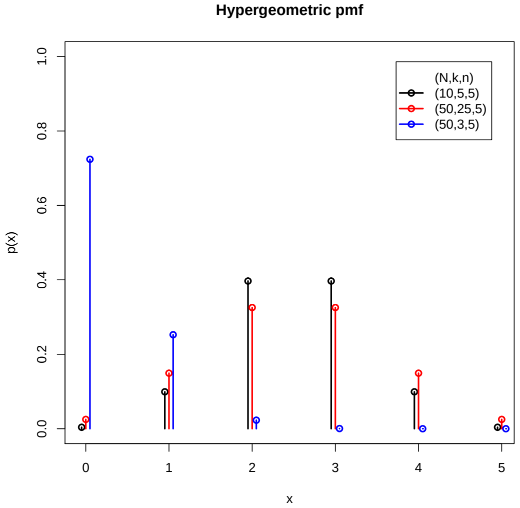
Hypergeometric distribution : properties

Mean and variance of the hypergeometric distribution

If $X \sim \text{Hyp}(N, k, n)$,

$$\mu = \mathbb{E}(X) = n\frac{k}{N} \qquad \text{and} \qquad \sigma^2 = \mathbb{V}\text{ar}(X) = n\frac{k}{N}\left(1 - \frac{k}{N}\right)\left(\frac{N-n}{N-1}\right)$$

Hypergeometric probability mass function



Hypergeometric distribution : remarks

- Let $\pi = \frac{k}{N}$, the initial proportion of 'defectives' in the full lot

Then,

$$\mathbb{E}(X) = n\pi \qquad \text{and} \qquad \mathbb{V}\text{ar}(X) = n\pi(1-\pi)\frac{N-n}{N-1}$$

- Compared to the Binomial mean and variance, we see that the dependence in successive trials affects the variance but **not** the mean
- The trials that compromise the experiment would be independent if we did **sampling with replacement**, namely, if each unit selected for the sample is replaced before the next one is drawn
- Similarly, the trials would be independent if we were sampling from an infinite set ($N = \infty$), as the proportion of defectives would remain constant for every trial ($k/(N-1) \sim k/N$, $(N-n)/(N-1) \sim 1$)

$\rightsquigarrow$ Binomial random variable

$\rightsquigarrow$ when N is 'large', the Binomial distribution is a good approximation to the Hypergeometric distribution

Hypergeometric distribution : examples

Example

The components of a 6-component system are to be randomly chosen from a bin of 20 used components. The resulting system will be functional if at least 4 of its 6 components are in working condition.

If 15 of the 20 components in the bin are in working condition, what is the probability that the resulting system will be functional?

- Let X be number of working components chosen. Then $X \sim \text{Hyp}(20, 15, 6)$
- The probability that the system will be functional is

$$\mathbb{P}(X \geq 4) = \mathbb{P}(X = 4) + \mathbb{P}(X = 5) + \mathbb{P}(X = 6) = \frac{\binom{15}{4}\binom{5}{2} + \binom{15}{5}\binom{5}{1} + \binom{15}{6}\binom{5}{0}}{\binom{20}{6}}$$

$$\simeq 0.8687$$

Hypergeometric distribution : examples

Example

The components of a 6-component system are to be randomly chosen from a bin of 20 used components. The resulting system will be functional if at least 4 of its 6 components are in working condition.

If 15 of the 20 components in the bin are in working condition, what is the probability that the resulting system will be functional?

- Let X be number of working components chosen. Then $X \sim \text{Hyp}(20, 15, 6)$
- The probability that the system will be functional is

$$\mathbb{P}(X \geq 4) = \mathbb{P}(X = 4) + \mathbb{P}(X = 5) + \mathbb{P}(X = 6) = \frac{\binom{15}{4}\binom{5}{2} + \binom{15}{5}\binom{5}{1} + \binom{15}{6}\binom{5}{0}}{\binom{20}{6}}$$

$$\simeq 0.8687$$

The Poisson distribution

The Poisson distribution

Assume you are interested in the number of occurrences of some random phenomenon in a fixed period of time

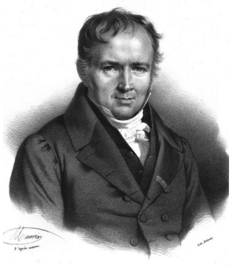
Define $X =$ number of occurrences. We say that X is a Poisson random variable with parameter λ, i.e.

$$X \sim \mathcal{P}(\lambda) \quad \text{or } X \sim \text{Pois}(\lambda),$$

if

$$p_{Pois}(x) = e^{-\lambda} \frac{\lambda^x}{x!} \qquad \text{for } x \in S_X = \{0, 1, 2, \dots\}$$

Note : Simeon-Denis Poisson (1781-1840) was a French mathematician



Poisson distribution : how does it arise?

- think of the time period of interest as being split up into a large number, say n, of subperiods
- assume that the phenomenon could occur at most one time in each of those subperiods, with some common probability π
- if what happens within one interval is independent to others,

$$X \sim \text{Bin}(n, \pi)$$

- now, as n increases, π should decrease (the shorter the period, the less likely the occurrence of the phenomenon) $\rightsquigarrow$ let $\pi = \lambda/n$ for some $\lambda > 0$
- then, for any $x \in \{0, 1, \dots, n\}$,

$$\begin{aligned}\mathbb{P}(X = x) &= \frac{n!}{x!(n-x)!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x} \\ &= \frac{n!}{n^x(n-x)!(1-\lambda/n)^x}\frac{\lambda^x}{x!}\left(1-\frac{\lambda}{n}\right)^n\end{aligned}$$

Poisson distribution : how does it arise?

- finally, as

$$\frac{n!}{n^x(n-x)!(1-\lambda/n)^x} \to 1 \quad \text{and} \quad \left(1-\frac{\lambda}{n}\right)^n \to e^{-\lambda}$$

as $n \to \infty$, it remains

$$\mathbb{P}(X=x) = e^{-\lambda}\frac{\lambda^x}{x!} \qquad \text{for } x \in \{0,1,\dots\}$$

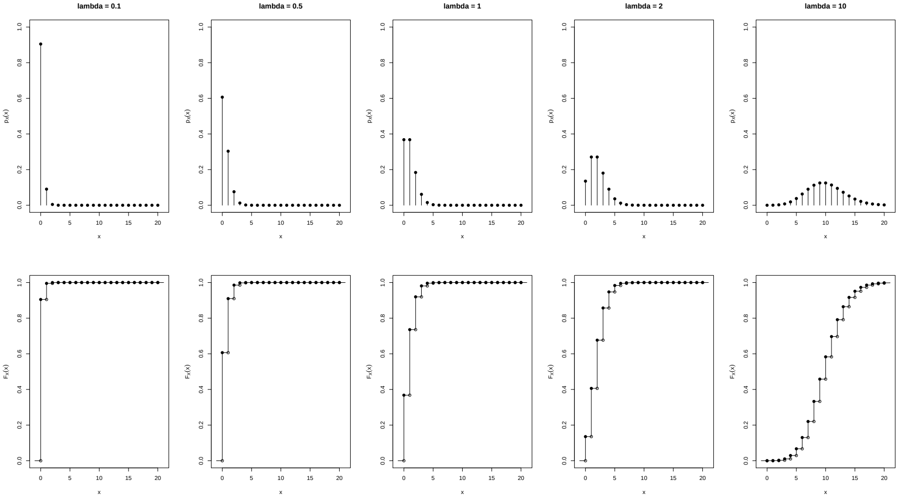
which is the Poisson distribution

- the Poisson distribution is thus suitable for modelling the number of occurrences of a random phenomenon satisfying some assumptions of continuity, stationarity, independence and non-simultaneity
- λ is called the intensity of the phenomenon

Note : we defined the $\mathcal{P}(\lambda)$ distribution by partitioning a time period, however the same reasoning can be applied to any interval, area or volume

$\rightsquigarrow$ the number of flaws in a new windshield might also be Poisson distributed

Poisson distribution : pmf and cdf



Poisson pmf and cdf, for $\lambda = \{0.1, 0.5, 1, 2, 10\}$

Poisson distribution : properties

- First we have

$$\sum_{x \in S_X} p(x) = \sum_{x=0}^{\infty} e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda} e^{\lambda} = 1$$

- Similarly,

$$\mathbb{E}(X) = \sum_{x \in S_X} x p(x) = \sum_{x=0}^{\infty} x e^{-\lambda} \frac{\lambda^x}{x!} = \lambda \sum_{x=1}^{\infty} e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!} = \lambda$$

$$\mathbb{E}(X^2) = \sum_{x \in S_X} x^2 p(x) = \ldots = \lambda^2 + \lambda$$

$\rightsquigarrow \mathbb{V}\mathrm{ar}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$

Poisson distribution : properties

Mean and variance of the Poisson distribution

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

$$\mathbb{E}(X) = \lambda \qquad \text{and} \qquad \mathbb{V}\text{ar}(X) = \lambda$$

Example

Over a 10-minute period, a counter records an average of 1.3 gamma particles per millisecond coming from a radioactive substance. To a good approximation, the distribution of the count, X, of gamma particles during the next millisecond is Poisson distributed. **Determine:**

a) λ,
b) the probability of oberving one or more gamma particles during the next millisecond and
c) the variance of this number

a) The mean of the Poisson distribution is λ, so we can approximate λ by the long-run average of the number of particles per millisecond, that is, $\lambda \simeq 1.3$. So we have

$$X \sim \mathcal{P}(1.3)$$

b) Thus,

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X = 0) = 1 - e^{-1.3}\frac{1.3^0}{0!} = 0.727$$

c) The variance of the Poisson distribution is also equal to λ, hence

$$\mathbb{V}\text{ar}(X) = 1.3 \;\; (\text{particles}^2)$$

Example

Over a 10-minute period, a counter records an average of 1.3 gamma particles per millisecond coming from a radioactive substance. To a good approximation, the distribution of the count, X, of gamma particles during the next millisecond is Poisson distributed. **Determine:**

a) λ,
b) the probability of oberving one or more gamma particles during the next millisecond and
c) the variance of this number

a) The mean of the Poisson distribution is λ, so we can approximate λ by the long-run average of the number of particles per millisecond, that is, $\lambda \simeq 1.3$. So we have

$$X \sim \mathcal{P}(1.3)$$

b) Thus,

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X = 0) = 1 - e^{-1.3}\frac{1.3^0}{0!} = 0.727$$

c) The variance of the Poisson distribution is also equal to λ, hence

$$\mathbb{V}\mathrm{ar}(X) = 1.3 \ \ (\text{particles}^2)$$

Poisson distribution : examples

Example

Suppose that the number of drivers who travel between a particular origin and destination during a designated time period has Poisson distribution with parameter $\lambda = 20$.
In the long-run, in what proportion of time periods will the number of drivers a) be at most 10? b) exceed 20? c) be between 10 and 20, inclusive? Strictly between 10 and 20?

Let X be the number of drivers. It is given that $X \sim \mathcal{P}(20)$

a)

$$\begin{aligned}\mathbb{P}(X \leq 10) &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \ldots + \mathbb{P}(X = 10) \\ &= e^{-20} + e^{-20} \times 20 + e^{-20}\frac{20^2}{2} + \ldots + e^{-20}\frac{20^{10}}{10!} \\ &= \ldots\end{aligned}$$

$\rightsquigarrow$ tedious ! $\rightsquigarrow$ use **statistical tables** or a statistical software

Poisson distribution : examples

Example

Suppose that the number of drivers who travel between a particular origin and destination during a designated time period has Poisson distribution with parameter $\lambda = 20$.
In the long-run, in what proportion of time periods will the number of drivers a) be at most 10? b) exceed 20? c) be between 10 and 20, inclusive? Strictly between 10 and 20?

Let X be the number of drivers. It is given that $X \sim \mathcal{P}(20)$

a)

$$\begin{aligned}\mathbb{P}(X \leq 10) &= \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \ldots + \mathbb{P}(X = 10) \\ &= e^{-20} + e^{-20} \times 20 + e^{-20}\frac{20^2}{2} + \ldots + e^{-20}\frac{20^{10}}{10!} \\ &= \ldots\end{aligned}$$

$\rightsquigarrow$ tedious ! $\rightsquigarrow$ use **statistical tables** or a statistical software

Poisson distribution : examples

Command Window

New to MATLAB? Watch this Video, see Demos, or read Getting Started.

```
>> poisscdf(10,20)

ans =

    0.0108

>> 1-poisscdf(20,20)

ans =

    0.4409

>> poisscdf(20,20)-poisscdf(9,20)

ans =

    0.5541

>> poisscdf(19,20)-poisscdf(10,20)

ans =

    0.4594

>> 
```

Poisson approximation to the Binomial distribution

- Since it was derived as a limit case of the Binomial distribution when n is 'large' and π is 'small', one can expect the Poisson distribution to be a good approximation to Bin(n, π) in that case $\rightsquigarrow$ true
- As it involves only one parameter, the Poisson pmf or the Poisson tables are usually easier to handle than the corresponding Binomial pmf and tables

Poisson approximation to the Binomial distribution

Example

It is known that 1% of the books at a certain bindery have defective bindings.

Compare the probabilities that x ($x = 0, 1, 2, \ldots$) of 100 books will have defective bindings using the (exact) formula for the binomial distribution and its Poisson approximation

- The exact Binomial pmf is $p_{Bin}(x) = \binom{100}{x} \times 0.01^x \times 0.99^{100-x}$, while its Poisson approximation is

$$p_{Pois}(x) = e^{-\lambda} \frac{\lambda^x}{x!}$$

with $\lambda = n \times \pi = 100 \times 0.01 = 1$

Poisson distribution : examples

Matlab computations give :

```
Command Window
New to MATLAB? Watch this Video, see Demos, or read Getting Started.

>> x=[0:100];
Bino=binopdf(x,100,0.01);
Poiss=poisspdf(x,1);
A=[x;Bino;Poiss];
A(:,1:8)

ans =

         0    1.0000    2.0000    3.0000    4.0000    5.0000    6.0000    7.0000
    0.3660    0.3697    0.1849    0.0610    0.0149    0.0029    0.0005    0.0001
    0.3679    0.3679    0.1839    0.0613    0.0153    0.0031    0.0005    0.0001

>>
```

We see that the error we would make by using the Poisson approximation instead of the true distribution is only of order 10^{-3}

$\leadsto$ very good approximation

The Uniform distribution

The Uniform distribution

There are also numerous **continuous** distributions which are of great interest. The simplest one is certainly the uniform distribution

- A random variable is said to be uniformly distributed over a **finite** interval $[\alpha, \beta]$, i.e.

$$\boxed{X \sim U_{[\alpha,\beta]}}$$

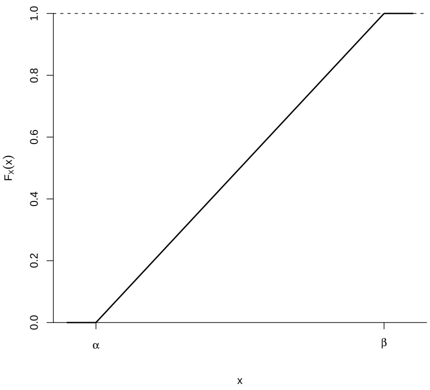
if its probability density function is given by

$$f(x) = \begin{cases} \frac{1}{\beta-\alpha} & \text{if } x \in [\alpha, \beta] \\ 0 & \text{otherwise} \end{cases} \qquad (\rightsquigarrow S_X = [\alpha, \beta])$$

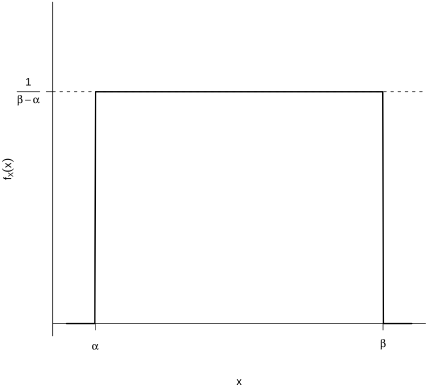
- Constant density $\rightsquigarrow X$ is just as likely to be “close” to any value in S_X
- By integration, it is easy to show that

$$F(x) = \begin{cases} 0 & \text{if } x < \alpha \\ \frac{x-\alpha}{\beta-\alpha} & \text{if } \alpha \leq x \leq \beta \\ 1 & \text{if } x > \beta \end{cases}$$

The Uniform distribution



cdf $F(x)$



pdf $f(x) = F'(x)$

Uniform distribution : properties

- Note that the constant Uniform density is set to $1/(\beta-\alpha)$ on $[\alpha,\beta]$ so as to ensure that $\int_\alpha^\beta f(x)\,dx = 1$
- Now,

$$\mathbb{E}(X) = \int_\alpha^\beta x\,\frac{1}{\beta-\alpha}dx = \frac{1}{\beta-\alpha}\left[\frac{x^2}{2}\right]_\alpha^\beta = \frac{\beta^2-\alpha^2}{2(\beta-\alpha)} = \frac{\alpha+\beta}{2}$$

- Similarly,

$$\mathbb{E}(X^2) = \int_\alpha^\beta x^2\frac{1}{\beta-\alpha}\,dx = \frac{1}{\beta-\alpha}\left[\frac{x^3}{3}\right]_\alpha^\beta = \frac{\beta^3-\alpha^3}{3(\beta-\alpha)} = \frac{\beta^2+\alpha\beta+\alpha^2}{3}$$

which implies $\mathbb{V}\mathrm{ar}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \ldots = \frac{(\beta-\alpha)^2}{12}$

Uniform distribution : properties

Mean and variance of the Uniform distribution

If $X \sim U_{[\alpha,\beta]}$,

$$\mathbb{E}(X) = \frac{\alpha+\beta}{2} \qquad \text{and} \qquad \mathbb{V}\text{ar}(X) = \frac{(\beta-\alpha)^2}{12}$$

Uniform distribution : example

Note:

The probability that X lies in any subinterval $[a, b]$ of $[\alpha, \beta]$ is simply equal to the length of that interval, i.e. $b - a$, divided by the length of the interval $[\alpha, \beta]$, i.e. $\beta - \alpha$:

$$\boxed{\mathbb{P}(a < X < b) = \frac{b-a}{\beta-\alpha}}$$

.

Uniform distribution : example

Example

Buses arrive at a specified stop at 15-minute intervals starting at 7 A.M. That is, they arrive at 7, 7:15, 7:30, 7:45, etc.

If a passenger arrives at the stop at a time this uniformly distributed between 7 and 7:30, find the probability that he waits less than 5 minutes for a bus

- Let X denote the time (in minutes) past 7 A.M. that the passenger arrives at the stop. We have $X \sim U_{[0,30]}$
- The passenger will have to wait less than 5 min if he arrives between 7:10 and 7:15 or between 7:25 and 7:30. This happens with probability

$$\mathbb{P}((10 < X < 15) \cup (25 < X < 30)) = \mathbb{P}(10 < X < 15) + \mathbb{P}(25 < X < 30)$$
$$= \frac{5}{30} + \frac{5}{30} = \frac{1}{3}$$

Uniform distribution : example

Example

Buses arrive at a specified stop at 15-minute intervals starting at 7 A.M. That is, they arrive at 7, 7:15, 7:30, 7:45, etc.

If a passenger arrives at the stop at a time this uniformly distributed between 7 and 7:30, find the probability that he waits less than 5 minutes for a bus

- Let X denote the time (in minutes) past 7 A.M. that the passenger arrives at the stop. We have $X \sim U_{[0,30]}$

- The passenger will have to wait less than 5 min if he arrives between 7:10 and 7:15 or between 7:25 and 7:30. This happens with probability

$$\mathbb{P}((10 < X < 15) \cup (25 < X < 30)) = \mathbb{P}(10 < X < 15) + \mathbb{P}(25 < X < 30)$$
$$= \frac{5}{30} + \frac{5}{30} = \frac{1}{3}$$

The Exponential distribution

The Exponential distribution

- Remind that a Poisson distributed r.v. counts the number of occurrences of a given phenomenon over a unit period of time
- The (random) amount of time before the first occurrence of that phenomenon is often of interest as well
- If $N \sim \mathcal{P}(\lambda)$ denote the number of occurrences over a unit period of time, then it can be shown that the number of occurrences of the phenomenon by a time x, say N_x, is $\sim \mathcal{P}(\lambda x)$ ("Poisson process")
- Denote X the amount of time before the first occurrence
- This time will exceed x ($x \geq 0$) if and only if there have been no occurrences of the phenomenon by time x, that is, $N_x = 0$
- As $N_x \sim \mathcal{P}(\lambda x)$, it follows

$\mathbb{P}(X > x) = \mathbb{P}(N_x = 0) = e^{-\lambda x}\frac{(\lambda x)^0}{0!} = e^{-\lambda x}$, which yields the cdf of X :

$$F(x) = 1 - e^{-\lambda x} \qquad \text{for } x \geq 0$$

This particular distribution is called the Exponential distribution

The Exponential distribution

- A random variable is said to be an Exponential random variable with parameter λ ($\lambda > 0$), i.e.

$$X \sim \text{Exp}(\lambda),$$

if its probability density function is given by

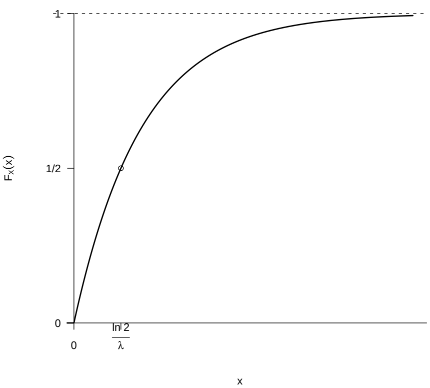
$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \qquad (\rightsquigarrow S_X = \mathbb{R}^+)$$

- By integration, it is easy to show that

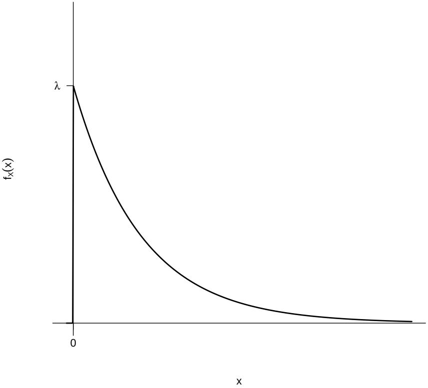
$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 - e^{-\lambda x} & \text{if } x \geq 0 \end{cases}$$

- From the above argument, it is easy to understand why this distribution is often useful for representing random amounts of time, like the amount of time required to complete a specified task, the waiting time at a counter, the amount of time until you receive a phone call, the amount of time until an earthquake occurs, etc.

The Exponential distribution



cdf $F(x)$



pdf $f(x) = F'(x)$

Exponential distribution : properties

- One can check that

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_0^{+\infty} \lambda e^{-\lambda x}\,dx = \lambda \left[\frac{e^{-\lambda x}}{(-\lambda)} \right]_0^{+\infty} = 1$$

- Moreover,

$$\mathbb{E}(X) = \int_0^{+\infty} x\,\lambda e^{-\lambda x} dx = \left[-x\, e^{-\lambda x} \right]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x} dx \qquad \text{(by Int.-by-parts)}$$

$$= 0 + \left[-\frac{e^{-\lambda x}}{\lambda} \right]_0^{+\infty} = \frac{1}{\lambda}$$

- Similarly, $\mathbb{E}(X^2) = \int_0^{+\infty} x^2 e^{-\lambda x}\,dx = \ldots = \frac{2}{\lambda^2}$, so that $\mathbb{V}\mathrm{ar}(X) = \frac{1}{\lambda^2}$.

Exponential distribution : properties

Mean and variance of the Exponential distribution

If $X \sim \text{Exp}(\lambda)$,

$$\mathbb{E}(X) = \frac{1}{\lambda} \qquad \text{and} \qquad \mathbb{V}\text{ar}(X) = \frac{1}{\lambda^2}$$

Exponential distribution : example

Example

Suppose that, on the average, 3 trucks arrive per hour to be unloaded at a warehouse.

What is the probability that the time between the arrivals of two successive trucks will be a) less than 5 minutes? b) at least 45 minutes?

- Assuming the number of trucks arriving during one hour is Poisson distributed (with parameter $\lambda = 3$). Then the amount of time X between two truck arrivals follows the Exp(3) distribution

- Hence,

a) $\mathbb{P}(X \leq 1/12) = \int_0^{1/12} 3e^{-3x}\,dx = 1 - e^{-1/4} = 0.221$

b) $\mathbb{P}(X > 3/4) = \int_{3/4}^{\infty} = e^{-9/4} = 0.105$

Exponential distribution : example

Example

Suppose that, on the average, 3 trucks arrive per hour to be unloaded at a warehouse.

What is the probability that the time between the arrivals of two successive trucks will be a) less than 5 minutes? b) at least 45 minutes?

- Assuming the number of trucks arriving during one hour is Poisson distributed (with parameter $\lambda = 3$). Then the amount of time X between two truck arrivals follows the Exp(3) distribution

- Hence,

a) $\mathbb{P}(X \leq 1/12) = \int_0^{1/12} 3e^{-3x}\, dx = 1 - e^{-1/4} = 0.221$

b) $\mathbb{P}(X > 3/4) = \int_{3/4}^{\infty} = e^{-9/4} = 0.105$

Other useful distributions

In the remainder of this we will also encounter some other continuous distributions, among these are

- the Student-*t* distribution, $X \sim t_\nu$;
- the χ^2 distribution, $X \sim \chi^2_\nu$;
- the Fisher (or just F) distribution, $X \sim F_{d_1,d_2}$

We will return to them later when we will need them

- The several distributions that we have introduced so far are very useful in the application of statistics to problems of engineering and physical science

-However, the most important distribution is certainly the

Normal distribution

which we will introduce in the next subchapter

Objectives

Now you should be able to :

- Understand the assumptions for some common discrete probability distributions □
- Select an appropriate discrete probability distribution to calculate probabilities in specific applications □
- Calculate probabilities, determine means and variances for some common discrete probability distributions □
- Understand the assumptions for some common continuous probability distributions □
- Select an appropriate continuous probability distribution to calculate probabilities in specific applications □
- Calculate probabilities, determine means and variances for some common continuous probability distributions □

Put yourself to the test ! $\rightsquigarrow$ Q9 p.128, Q13 p.130, Q20 p.131, Q33 p.133