Mathematics for Data Science

Lecture 2: Continuous Random Variables

Mohamad GHASSANY

EFREI PARIS

Standard Normal distribution $\mathcal{N}(0,1)$

Relation between Normal distribution and Standard Normal distribution

Distribution function of Standard Normal distribution $\mathcal{N}(0,1)$

Normal Approximation to the Binomial Distribution

Other related distributions

Continuous Random Variable

- Previously we have dealt with Discrete Random Variables, i.e. variables whose universe is finite or countable.
- There are however variables whose universe is infinite uncountable.
- Examples:
  - The arrival time of a train at a given station.
  - The lifetime of a transistor.

Definition

*X is a continuous random variable [1] **with density** if there exists a non-negative function f defined for any $x \in \mathbb{R}$ and verifying for any set B of real numbers the property*

$$P(X \in B) = \int_B f(x)dx$$

The function f is called density function of the random variable X.

- All probability questions related to X can be treated with f.
- For example if $B = [a, b]$, we get:

$$P(a \leqslant X \leqslant b) = \int_a^b f(x)dx$$

[1] Not all Continuous Random Variable have a density function.

Graphically, $P(a \leqslant X \leqslant b)$ is the area of the surface between the x-axis, the curve corresponding to $f(x)$ and the lines $x = a$ and $x = b$.

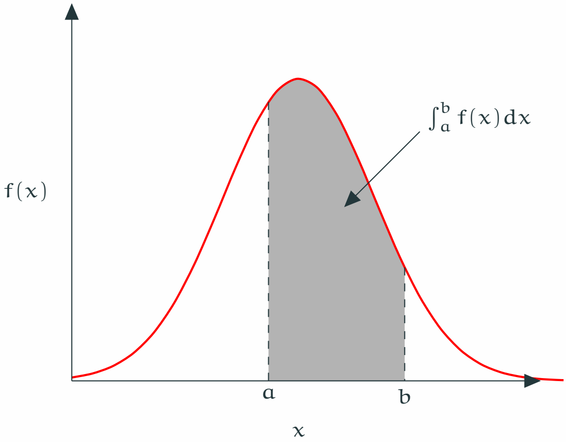


Figure 1: $P(a \leqslant X \leqslant b) =$ *area of shaded surface*

Continuous Random Variable - Graphical Interpretation - Example

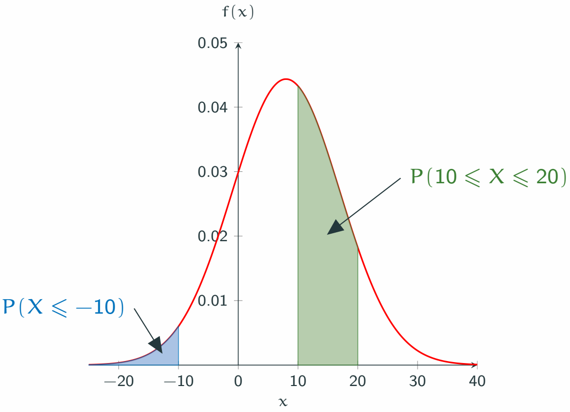
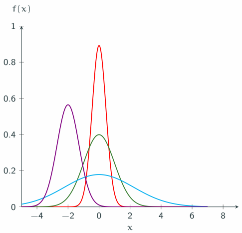
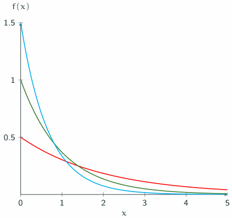
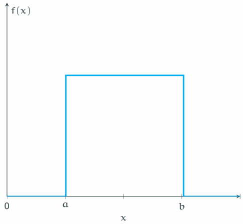


Figure 2: The colored areas corresponds to probabilities. $f(x)$ *being a probability density function.*







Properties

For any continuous random variable X of density f:

- $f(x) \geqslant 0 \quad \forall x \in \mathbb{R}$
- $\int_{-\infty}^{+\infty} f(x)dx = 1$
- Since $P(a \leqslant X \leqslant b) = \int_a^b f(x)dx$, if $a = b$ then $P(X = a) = \int_a^a f(x)dx = 0$
- This means that the probability of a continuous random variable taking a fixed isolated value is always zero.

Continuous Random Variable - Examples

Example

Let X be the random real variable of probability density

$$f(x) = \begin{cases} kx & \text{if} \quad 0 \leqslant x \leqslant 5 \\ 0 & \text{if not} \end{cases}$$

1. Calculate k.
2. Calculate: $P(1 \leqslant X \leqslant 3), P(2 \leqslant X \leqslant 4)$ and $P(X < 3)$.

Example

Let X be a continuous random variable with density function

$$f(x) = \begin{cases} \frac{1}{6}x + k & \text{if} \quad 0 \leqslant x \leqslant 3 \\ 0 & \text{if not} \end{cases}$$

1. Calculate k.
2. Calculate $P(1 \leqslant X \leqslant 2)$

Distribution function of continuous random variables

Distribution function of continuous random variables

Definition

If as for Random Variable Discrete, we define the distribution function of X by:

$$\begin{aligned} F_X : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto F_X(a) = P(X \leqslant a) \end{aligned}$$

then the relation between the distribution function F_X and the probability density function $f(x)$ is the following:

$$\forall \quad a \in \mathbb{R} \quad F_X(a) = P(X \leqslant a) = \int_{-\infty}^{a} f(x)dx$$

Properties

For a continuous random variable X:

- $F'_X(x) = \frac{d}{dx}F_X(x) = f(x)$.
- For all real numbers $a \leqslant b$,

$$\begin{aligned} P(a < X < b) = P(a < X \leqslant b) &= P(a \leqslant X < b) \\ &= P(a \leqslant X \leqslant b) = F_X(b) - F_X(a) = \int_a^b f(x)dx \end{aligned}$$

The distribution function corresponds to the cumulative probabilities associated with the continuous random variable on an interval.

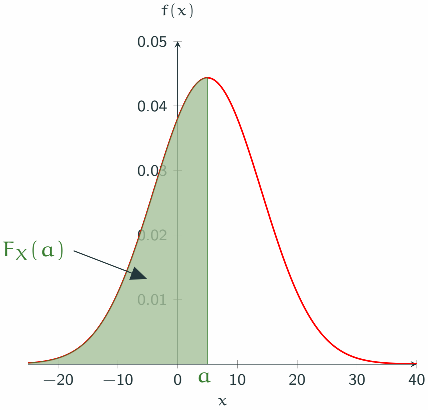


Figure 3: The area shaded in green under the curve of the density function corresponds to the probability $P(X < a) = F_X(a)$ *and is* 0.5 *because this corresponds exactly to half of the total area under the curve.*

Properties

The properties of the distribution function are as follows:

1. F_X is continuous on $\mathbb{R}$, derivable at any point where f is continuous.
2. F_X is increasing on $\mathbb{R}$.
3. F_X has values in $[0, 1]$.
4. $\lim\limits_{x \to -\infty} F_X(x) = 0$ and $\lim\limits_{x \to +\infty} F_X(x) = 1$.

Example

Let X and Y two random variables of density functions:

$$f_X(x) = \begin{cases} kx & \text{if} \quad 0 \leqslant x \leqslant 5 \\ 0 & \text{if not} \end{cases}$$

and

$$f_Y(y) = \begin{cases} \frac{1}{6}y + k & \text{if} \quad 0 \leqslant y \leqslant 3 \\ 0 & \text{if not} \end{cases}$$

Calculate $F_X(a)$ and $F_Y(a)$ for all $a \in \mathbb{R}$.

Function of a continuous random variable

Function of a continuous random variable

- Let X be a continuous random variable with density f_X and distribution function F_X.
- Let h be a continuous function defined on $X(\Omega)$, then $Y = h(X)$ is a random variable.
- To determine the density of Y, denoted f_Y, we first compute the distribution function of Y, denoted F_Y, then we derivate it to determine f_Y.

Calculating the densities

Let X be a continuous random variable with density f_X and distribution function F_X. Find the density function of the following random variables:

- $Y = aX + b$
- $Z = X^2$
- $T = e^X$

Example

Let X a random variable having the density function:

$$f_X(x) = 2x \times \mathbb{1}_{[0,1]}(x)$$

Determine the density function of: $Y = 3X + 1$, $Z = X^2$ and $T = e^X$.

Moments of Continuous Random Variable

Expected value

Definition

If X is a continuous random variable of density f, we call the expected value of X, the real $E(X)$, defined by:

$$E(X) = \int_{-\infty}^{+\infty} x f(x)\,dx$$

if it exists.

The properties of the expected value of a continuous random variable are the same as for a discrete random variable.

Properties

Let X be a continuous random variable,

- $E(aX + b) = aE(X) + b \qquad a \geqslant 0$ and $b \in \mathbb{R}$.
- If $X \geqslant 0$ then $E(X) \geqslant 0$.
- If X and Y are two Random Variables defined on the same universe Ω then

$$E(X + Y) = E(X) + E(Y)$$

Transfer theorem

Theorem

If X *is a random variable of density* $f(x)$*, then for any real function* g *we have*

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)dx$$

Example

Let X a random variable of density

$$f_X(x) = \begin{cases} 2x & \text{if} \quad 0 \leqslant x \leqslant 1 \\ 0 & \text{if not} \end{cases}$$

Calculate the expected value of $Y = 3X + 1$, $Z = X^2$ and $T = e^X$.

The variance of a random variable $V(X)$ is a dispersion parameter which corresponds to the centered moment of order 2 of the random variable X.

Definition

If X is a random variable with expectation $E(X)$, we call the variance of X the real

$$V(X) = E\left([X - E(X)]^2\right) = E(X^2) - [E(X)]^2$$

If X is a continuous random variable, we compute $E(X^2)$ using the transfer theorem,

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)\,dx$$

Example

Calculate la variance of X defined in the previous example.

Properties

If X is a random variable with a variance then:

- $V(X) \geqslant 0$, if it exists.
- $\forall\ a \in \mathbb{R}, V(aX) = a^2V(X)$
- $\forall\ (a, b) \in \mathbb{R}, V(aX + b) = a^2V(X)$
- If X and Y are two independent Random Variables, $V(X + Y) = V(X) + V(Y)$

Definition

If X is a random variable with variance $V(X)$, we call the standard deviation of X the real:

$$\sigma_X = \sqrt{V(X)}$$

Lois Usuelles de Variables Aléatoires Continues

Définition

La variable aléatoire X suit une loi uniforme sur le segment $[a,b]$ avec $a<b$ si sa densité de probabilité est donnée par

$$f(x)=\begin{cases}\dfrac{1}{b-a} & \text{si} \quad x\in[a,b]\\ 0 & \text{si} \quad x\notin[a,b]\end{cases} = \frac{1}{b-a}\mathbb{1}_{[a,b]}(x)$$

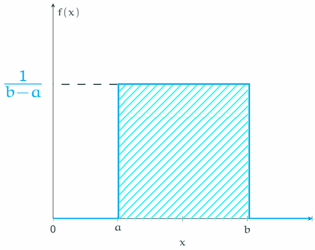


Figure 4: Fonction de densité de $\mathcal{U}([a,b])$

Uniform ditribution $U(a, b)$

Definition

The random variable X follows a Uniform distribution on the segment $[a, b]$ with $a < b$ if its density function is given by

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{if} \quad x \in [a, b] \\ 0 & \text{if} \quad x \notin [a, b] \end{cases} = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$$

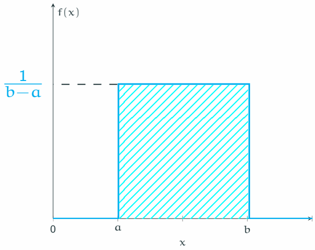


Figure 5: Density function of $\mathcal{U}([a, b])$

Uniform distribution $\mathcal{U}(a,b)$

- The *distribution function* associated to the continuous uniform distribution is

$$F_X(x) = \begin{cases} 0 & \text{if} \quad x < a \\ \dfrac{x-a}{b-a} & \text{if} \quad a \leqslant x \leqslant b \\ 1 & \text{if} \quad x > b \end{cases}$$

- $E(X) = \dfrac{b+a}{2}$
- $V(X) = \dfrac{(b-a)^2}{12}$

Exponential distribution $\mathcal{E}(\lambda)$

Definition

A random variable X is **exponential** (or follows an Exponential distribution) of parameter λ if its density function is given by

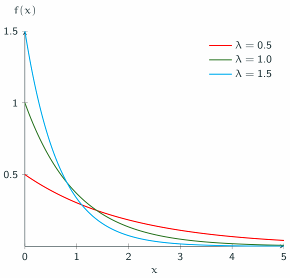
$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if} \quad x \geqslant 0 \\ 0 & \text{if} \quad x < 0 \end{cases} = \lambda e^{-\lambda x} \mathbb{1}_{\mathbb{R}^+}(x)$$

We say $X \sim \mathcal{E}(\lambda)$

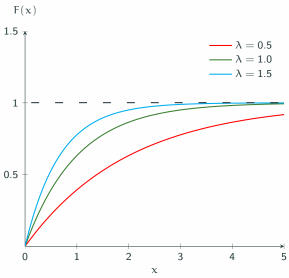
- The distribution function F of an exponential random variable is given by

$$\text{if } x \geqslant 0 \quad F(x) = P(X \leqslant x) = 1 - e^{-\lambda x}$$

- $E(X) = \dfrac{1}{\lambda}$ and $V(X) = \dfrac{1}{\lambda^2}$



(a) Density function of exponential distribution



(b) Distribution function of exponential distribution

Use cases of the exponential distribution:

- Represent the waiting time before the arrival of a specified event.
- To model the lifetime of a phenomenon without memory or without aging.
- A nonnegative random variable X is said to be *memoryless* when

$$P(X > t + h | X > t) = P(X > h) \qquad \forall \quad t, h \geqslant 0$$

- For example, the lifetime of radioactivity or of an electronic component.

Normal distribution or Gaussian

$\mathcal{N}(\mu, \sigma^2)$

Normal distribution $\mathcal{N}(\mu, \sigma^2)$

Definition

A random variable X is said to be **normal** (or Gaussian) with parameters μ and σ^2 if its density function is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \qquad \forall\, x \in \mathbb{R}$$

With $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$. We say that $X \sim \mathcal{N}(\mu, \sigma^2)$.

Moments of Normal distribution $\mathcal{N}(\mu, \sigma^2)$

- $E(X) = \mu$
- $V(X) = \sigma^2$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \qquad \forall\, x \in \mathbb{R}$$

- The function f is **even** with axis of symmetry $x = \mu$ car $f(x+\mu) = f(\mu - x)$.
- $f'(x) = 0$ when $x = \mu$, $f'(x) < 0$ when $x > \mu$ and $f'(x) > 0$ when $x < \mu$

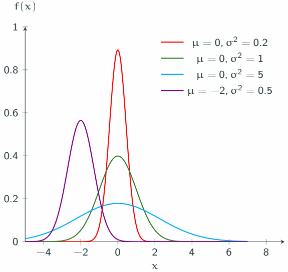


Figure 7: ***Remark:*** *The parameter μ represents the axis of symmetry and σ the degree of flatness of the curve of the Normal distribution whose shape is that of a bell curve.*

Stability of the Normal distribution

Theorem

- $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$
- $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$
- X_1 et X_2 are independent.

So $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

Standard Normal distribution $\mathcal{N}(0, 1)$

Standard Normal distribution $\mathcal{N}(0,1)$

Definition

A continuous random variable X follows a Standard Normal distribution if its density function is given by

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \qquad \forall\, x \in \mathbb{R}$$

On dit $X \sim \mathcal{N}(0,1)$.

Moments of Standard Normal distribution $\mathcal{N}(0,1)$

- $E(X) = 0$
- $V(X) = 1$

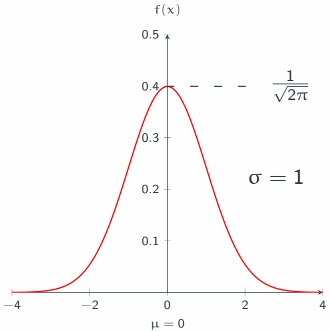


Figure 8: Density of Standard Normal distribution $\mathcal{N}(0, 1)$.

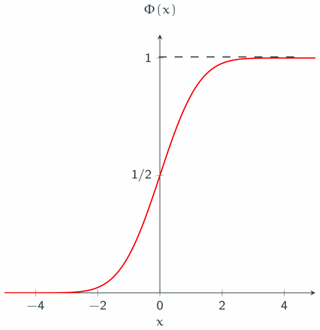


Figure 9: Distribution function of Standard Normal distribution $\mathcal{N}(0, 1)$.

Relation between Normal distribution and Standard Normal distribution

Theorem

If X follows a Normal distribution $\mathcal{N}(\mu, \sigma^2)$, then $Z = \dfrac{X - \mu}{\sigma}$ is a random variable that follows the Standard Normal distribution $\mathcal{N}(0, 1)$.

Distribution function of Standard Normal distribution $\mathcal{N}(0, 1)$

Distribution function of Standard Normal distribution $\mathcal{N}(0, 1)$

The distribution function of Standard Normal distribution allows to obtain the probabilities associated to all normal random variables $\mathcal{N}(\mu, \sigma^2)$ after transformation into a standardised variable.

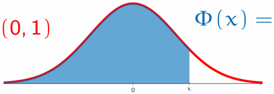
Definition

Let us call Φ, the distribution function of Standard Normal distribution $\mathcal{N}(0, 1)$, such that

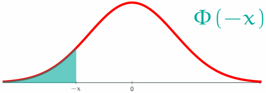
$$\forall\, x \in \mathbb{R} \quad \Phi(x) = P(X \leqslant x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} f(t)dt$$

Distribution function of Standard Normal distribution $\mathcal{N}(0,1)$

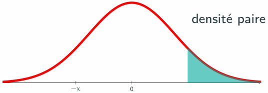
$X \sim \mathcal{N}(0,1)$



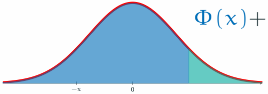
$\Phi(x) = P(X \leqslant x)$



$\Phi(-x) = P(X \leqslant -x)$



densité paire $\Rightarrow$ courbe symétrique % $x = 0$



$\Phi(x) + \Phi(-x) = 1$

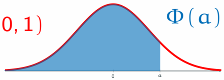
Properties de Φ

The properties associated with the distribution function Φ are:

1. Φ is increasing, continuous and derivable on $\mathbb{R}$ and verifies: $\lim\limits_{x \to -\infty} \Phi(x) = 0$ and $\lim\limits_{x \to \infty} \Phi(x) = 1$
2. $\forall\, x \in \mathbb{R} \quad \Phi(x) + \Phi(-x) = 1$
3. $\forall\, x \in \mathbb{R} \quad \Phi(x) - \Phi(-x) = 2\Phi(x) - 1$

Standard Normal distribution table $\mathcal{N}(0,1)$

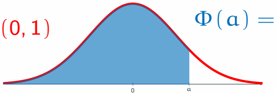
$X \sim \mathcal{N}(0,1)$



$\Phi(a) = P(X \leqslant a)$

a	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

Standard Normal distribution table $\mathcal{N}(0,1)$



a	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

For example, for $x = 1.23$ (intersection of the line 1.2 and the column 0.03), we get: $\Phi(1.23) \approx$ 0.8907.

Calculation of Normal distribution probabilities

Exemple 1

Let X be a random variable with standard normal distribution. Calculate:

1. $P(X > 2)$
2. $P(2 < X < 5)$

Exemple 2

SLet X be a random variable with normal distribution of parameters $\mu = 3$ et $\sigma^2 = 4$. Calculate:

1. $P(X > 0)$
2. $P(2 < X < 5)$
3. $P(|X - 3| > 4)$

Normal Approximation to the Binomial Distribution

Moivre Laplace Theorem

Suppose that for all n, X_n follows a binomial distribution $\mathcal{B}(n, p)$ with $p \in]0, 1[$.

Then the variable $Z_n = \dfrac{X_n - np}{\sqrt{np(1-p)}}$ converges in law to a Standard Normal distribution $\mathcal{N}(0, 1)$.

- This result was progressively generalized by Laplace, Gauss and others to become the Theorem currently known as the Central Limit Theorem which is one of the two most important results in probability theory.
- In practice, many random phenomena follow approximately a normal distribution.

Normal Approximation to the Binomial Distribution

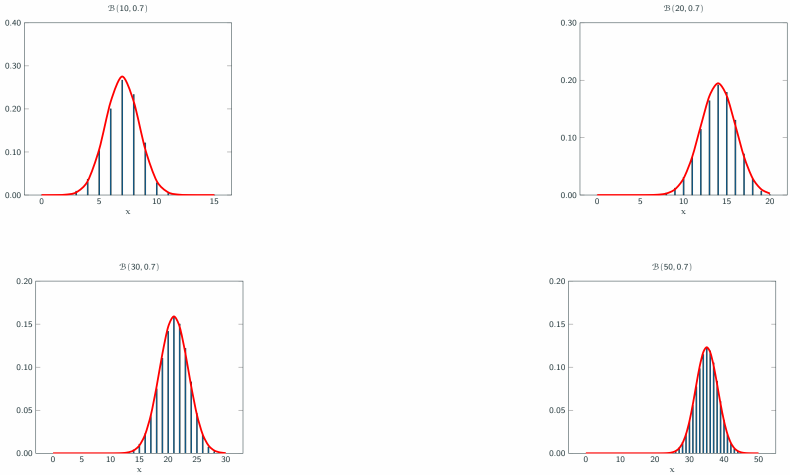


Figure 10: The probability distribution of a random variable $\mathcal{B}(n, p)$ becomes more and more "normal" as n increases.

Other related distributions

Definition

Let $X_1, X_2, \ldots, X_n$ **standard normal** random variables, and Y the random variable defined by

$$Y = X_1^2 + X_2^2 + \ldots + X_i^2 + \ldots + X_n^2 = \sum_{i=1}^{n} X_i^2$$

We say that Y follows the χ^2 distribution (or Pearson's law) with n degrees of freedom, $Y \sim \chi^2(n)$

- The χ^2 distribution has many applications in the context of comparison of proportions, tests of conformity of an observed distribution to a theoretical distribution and the test of independence of two qualitative variables. These are the chi-square tests.
- Note: If $n = 1$, the χ^2 corresponds to the square of a standard normal variable $\mathcal{N}(0, 1)$.

Definition

Let U be a random variable following a **standard normal distribution** $\mathcal{N}(0,1)$ and V a random variable following a $\chi^2(n)$ distribution, U and V being independent, we say that $T_n = \dfrac{U}{\sqrt{\frac{V}{n}}}$ follows a Student's distribution with n degrees of freedom. $T_n \sim St(n)$

- Student distribution is used in tests of comparison of parameters (such as the mean) and in the estimation of population parameters from sample data (Student test).

Fisher-Snedecor distribution $\mathcal{F}(n, m)$

Definition

Let U and V be two independent random variables following a χ^2 distribution with n and m degrees of freedom respectively.

We say that $F = \dfrac{U/n}{V/m}$ follows a Fisher-Snedecor distribution with (n, m) degrees of freedom. $F \sim \mathcal{F}(n, m)$

- The Fisher-Snedecor distribution is used to compare two observed variances and is used especially in the numerous analysis of variance and covariance tests.

Extra content

- A visual introduction to probability and statistics: http://students.brown.edu/seeing-theory/
- Distribution Calculator: https://gallery.shinyapps.io/dist_calc/