5. Continuous Random Variables

CONTINUOUS RANDOM VARIABLE

- It is a rv whose range is an interval over the real line.
  - Weight of an item, time until failure of a mechanical component, length of an object, ...

random variable Y which denotes the amount of rainfall we get in a month.

Why do we treat continuous and discrete random variables differently?

Example Imagine that you are asked to build a metal sheet with thickness exactly as 0.3mm.
It is impossible ! Sure, you can get the thickness very very close to 0.3mm, like between 0.29999mm and 0.30001mm. But it is in theory impossible to get exactly 0.3mm since tiny tiny error always exist.

5.1 Introduction

The previous example tells us:

If X is a continuous random variable, then P(X=t)=0 for any number t.

There is no meaning to study P(X=t) for continuous R.V. Instead, we consider P(a<X<b) for interval (a, b)

$$P(a < X < b) \quad ? \quad P(a \leq X \leq b)$$

How do we study the probability of an interval ?

5.1 Introduction

Definitions *We say that X is a **continuous** random variable if there exists a nonnegative function f_X, defined for all real $x \in \mathbb{R}$, having the property that, for any set B of real numbers,*

$$P(X \in B) = \int_B f_X(x)\,\mathrm{dx}\,.$$

*The function f_X is called the **probability density function** (p.d.f.) of the random variable X.*

All probability statements about X can be answered in terms of f_X. For instance, letting $B = [a,b]$, we obtain

$$P(a \le X \le b) = \int_a^b f_X(x)\,\mathrm{dx}\,. \qquad (*)$$

5.1 Introduction

Example

$$P(X = a) = P(a \leq X \leq a) = \int_a^a f(x)dx = 0$$

Definitions *We define the **distribution function** of X by*

$$F_X(x) = P(X \leq x), \quad \textit{for } x \in \mathbb{R}.$$

Remark Note that the definition for distribution function is the same for discrete and continuous random variables. Therefore, in the context of continuous random variable,

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{dt}, \ \ x \in \mathbb{R}$$

5.1 Introduction

and, using the Fundamental Theorem of Calculus,

$$F_X'(x) = f_X(x), \;\; x \in \mathbb{R}.$$

That is, the density is the derivative of the cumulative distribution function. A more intuitive interpretation:

$$P\left(x - \frac{\varepsilon}{2} \leq X \leq x + \frac{\varepsilon}{2}\right) = \int_{x-\varepsilon/2}^{x+\varepsilon/2} f_X(x)\,\mathrm{dx} \approx \varepsilon f(x),$$

when ε is small and when $f(\cdot)$ is continuous at x.

The probability that X will be contained in an interval of length ε around the point x is approximately $\varepsilon f(x)$. From this result we see that $f(x)$ is a measure of how likely it is that the random variable will be near x.

5.1 Introduction

Remarks

$P(X = x) = 0$ for any $x \in (-\infty, \infty)$.

The distribution function, F_X, is continuous.

For any $a, b \in (-\infty, \infty)$,

$$\begin{aligned} P(a \leq X \leq b) &= P(a < X \leq b) \\ &= P(a \leq X < b) \\ &= P(a < X < b). \end{aligned}$$

5.1 Introduction

Determine the constant in the probability density function

$$1 = P(-\infty < X < \infty) = \int_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x\,.$$

Example Suppose that X is a continuous random variable whose probability density function is of the form:

$$f_X(x) = \begin{cases} c(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{otherwise}\,. \end{cases}$$

(i) What is the value of c? (ii) Compute $P(X > 1)$.

Solution

(i) $1 = \int_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x = c\int_0^2 (4x - 2x^2)\,\mathrm{d}x = \frac{8c}{3}$. Therefore $c = 3/8$.

(ii) $P(X > 1) = \int_1^{\infty} f_X(x)\,\mathrm{d}x = \frac{3}{8}\int_1^2 (4x - 2x^2)\,\mathrm{d}x = \frac{1}{2}$.

5.1 Introduction

Example An electrical appliance will function for a random amount of time T. If the probability density function of T is given in the form (where $\lambda >$ is a fixed constant):

$$f_T(t) = \begin{cases} ce^{-\lambda t} & 0 < t < \infty \\ 0 & \text{otherwise .} \end{cases}$$

(i) What is the value of c? (ii) Show that for any $s, t > 0$, then $P(T > s+t | T > s) = P(T > t)$.

Solution (i) $1 = \int_{-\infty}^{\infty} f_T(t)\,\mathrm{dt} = c\int_0^{\infty} e^{-\lambda t}\,\mathrm{dt} = \frac{c}{\lambda}.$ Hence $c = \lambda$.

(ii) First consider $P(T > t) = \int_t^{\infty} f_T(x)\,\mathrm{dx} = \lambda\int_t^{\infty} e^{-\lambda x}\,\mathrm{dx} = e^{-\lambda t}.$

Therefore
$$\begin{aligned} P(T > s+t | T > s) &= \frac{P(T > s+t; T > s)}{P(T > s)} \\ &= \frac{P(T > s+t)}{P(T > s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P(T > t). \end{aligned}$$

5.1 Introduction

Remarks

$$f_T(t) = \begin{cases} ce^{-\lambda t} & 0 < t < \infty \\ 0 & \text{otherwise} . \end{cases}$$

The distribution of T is commonly called a **exponential distribution** with parameter λ.

This particular property $P(T > s+t | T > s) = P(T > t).$

is called the **memoryless property**.

The cumulative distribution function is

$$F_T(t) = \begin{cases} 1 - e^{-\lambda t}, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

5.1 Introduction

Example The lifetime in hours of a certain kind of radio tube is a random variable having a probability density function given by

$$f(x) = \begin{cases} 0, & x \leq 100 \\ 100/x^2, & x > 100 \end{cases}.$$

What is the probability that exactly 2 of 5 such tubes in a radio will have to be replaced within the first 150 hours of operation? Assume that the radio tubes function independently.

For $1 \leq i \leq 5$, let E_i be the event that the ith radio tube has to be replaced within 150 hours of operation.
Let X be the number of radio tubes (out of 5) that have to be replaced within 150 hours of operation. Then $X \sim \mathrm{Bin}(5, p)$ where $p = P(E_1)$.

We are asked $P(X=2)$. This is given as $\binom{5}{2} p^2 (1-p)^3$.

Let T_i be the lifetime of the radio tube i. We need to compute

$$p = P(T_1 \leq 150) = \int_{-\infty}^{150} f(t)\,\mathrm{d}t = \int_{100}^{150} \frac{100}{t^2}\,\mathrm{d}t = \frac{1}{3}.$$

Probability asked is $\binom{5}{2}(1/3)^2(1-1/3)^3 = 80/243 = 0.3292.$

Solution

5.2 Expectation and Variance of Continuous Random Variables

Definitions *Let X be a continuous random variable with probability density function f_X, then*

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{dx} \quad \textit{and} \\ \mathrm{var}(X) &= \int_{-\infty}^{\infty} (x - E(X))^2 f_X(x)\,\mathrm{dx}\,. \end{aligned}$$

A number of results in the discrete case either carry over or with obvious change to the continuous case

5.2 Expectation and Variance of Continuous Random Variables

Theorem *If X is a continuous random variable with probability density function f_X, then for any real value function g*

(a)

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\,\mathrm{dx}\,.$$

(b) Same linearity Property:

$$E(aX+b) = aE(X)+b.$$

(c) Same alternative formula for variance:

$$\mathrm{var}(X) = E(X^2) - [E(X)]^2.$$

5.2 Expectation and Variance of Continuous Random Variables

Theorem (Tail sum formula). *Suppose X is a nonnegative continuous random variable, then*

$$E(X) = \int_0^\infty P(X > x)\,\mathrm{dx} = \int_0^\infty P(X \geq x)\,\mathrm{dx}\,.$$

Proof We have

$$\int_0^\infty P(X > x)\,\mathrm{dx} = \int_0^\infty \int_x^\infty f(y)\,\mathrm{dy}\,\mathrm{dx}$$

where we have used the fact that $P(X > x) = \int_x^\infty f(y)\,\mathrm{dy}$. Interchanging the order of integration we get

$$\int_0^\infty P(X > x)\,\mathrm{dx} = \int_0^\infty \left(\int_0^y \mathrm{dx} \right) f(y)\,\mathrm{dy} = \int_0^\infty y f(y)\,\mathrm{dy} = E[X].$$

5.2 Expectation and Variance of Continuous Random Variables

Example Find the mean and the variance of the random variable, X, with the probability density function given as

$$f(x) = \begin{cases} 1/(b-a), & a \leq x \leq b \\ 0, & \text{otherwise} \end{cases}.$$

Here $a < b$ are given constants. X is said to be a **uniform distribution** over $[a, b]$.

Solution

Mean of X:

$$E(X) = \int_{-\infty}^{\infty} x f(x)\,\mathrm{dx} = \int_a^b \frac{x}{b-a}\,\mathrm{dx} = \frac{a+b}{2}.$$

Variance of X:

$$\begin{aligned} \operatorname{var}(X) &= E(X^2) - \left(\frac{a+b}{2}\right)^2 \\ &= \int_a^b \frac{x^2}{b-a}\,\mathrm{dx} - \frac{(a+b)^2}{4} = \frac{b^3-a^3}{3(b-a)} - \frac{(a+b)^2}{4} \\ &= \frac{b^2+ab+a^2}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}. \end{aligned}$$

5.2 Expectation and Variance of Continuous Random Variables

Example Find $E\left(e^X\right)$ where X has probability density function given by

$$f(x) = \begin{cases} 2x, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}.$$

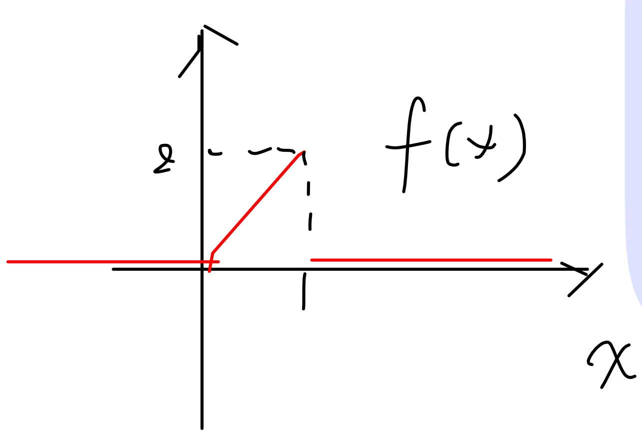
$$\int_{-\infty}^{+\infty} e^x f(x)\,dx = \int_0^1 2e^x x\,dx = 2\int_0^1 xe^x\,dx$$

Solution

$$E\left(e^X\right) = \int_0^1 e^x 2x\,\mathrm{dx} = 2\left[(x-1)e^x\right]_0^1 = 2.$$

cdf: $F(x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & \text{if } x < 0 \\ x^2 & 0 \le x \le 1 \\ 1, & 1 < x \end{cases}$

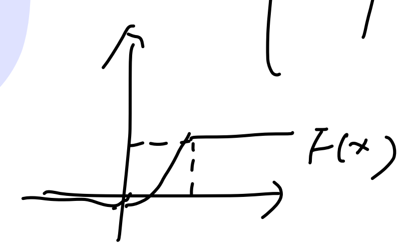
pdf:



Theorem We have

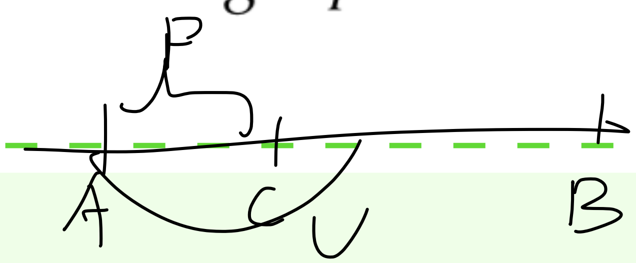
(a) $\mathrm{var}(aX+b) = a^2\,\mathrm{var}(X)$.

(b) $\mathrm{SD}(aX+b) = |a|\,\mathrm{SD}(X)$.



5.2 Expectation and Variance of Continuous Random Variables

Example A stick of length 1 is broken at random. Determine the expected length of the piece that contains the point of length p from one end, call A, where $0 \leq p \leq 1$.



Solution Let U be the point (measuring from A) where the stick is broken into two pieces. Then the probability density function of U is given by

$$f_U(u) = \begin{cases} 1, & \text{for } 0 < u < 1 \\ 0, & \text{otherwise} \end{cases}.$$

Let $L_p(U)$ denote the length of the piece which contains the point p, and note that

$$L_p(U) = \begin{cases} 1 - U, & \text{if } U < p \\ U, & \text{if } U > p \end{cases}.$$

Hence

$$\begin{aligned} E[L_p(U)] &= \int_{-\infty}^{\infty} L_p(u) \cdot f_U(u) du = \int_0^1 L_p(u) \cdot 1 du \\ &= \int_0^p (1-u) du + \int_p^1 u du = 1/2 + p(1-p). \end{aligned}$$

5.2 Expectation and Variance of Continuous Random Variables

Example Suppose that if you are s minutes early for an appointment, then you incur a cost cs, and if you are s minutes late, then you incur a cost ks. Suppose the travelling time from where you are to the location of your appointment is a continuous random variable having probability density function f. Determine the time at which you should depart if you want to minimize your expected cost.

Solution Let X be your travel time. If you leave t minutes before your appointment, then the cost, call it $C_t(X)$, is given by

$$C_t(X) = \begin{cases} c(t-X), & \text{if } X \leq t \\ k(X-t), & \text{if } X > t \end{cases}.$$

Therefore,

$$\begin{aligned} E[C_t(X)] &= \int_0^\infty C_t(x) f(x)\,\mathrm{d}x \\ &= \int_0^t c(t-x) f(x)\,\mathrm{d}x + \int_t^\infty k(x-t) f(x)\,\mathrm{d}x \\ &= ct \int_0^t f(x)\,\mathrm{d}x - c \int_0^t x f(x)\,\mathrm{d}x + k \int_t^\infty x f(x)\,\mathrm{d}x - kt \int_t^\infty f(x)\,\mathrm{d}x \\ &= ctF(t) - c \int_0^t x f(x)\,\mathrm{d}x + k \int_t^\infty x f(x)\,\mathrm{d}x - kt[1-F(t)]. \end{aligned}$$

why differentiate $= -ktf(t)$?

next slide

5.2 Expectation and Variance of Continuous Random Variables

Continue from last slide

Solution

The value of t which minimizes $E[C_t(X)]$ is obtained by calculus. Differentiation yields

$$\begin{aligned}\frac{d}{\mathrm{dt}}E[C_t(X)] &= cF(t)+ctf(t)-ctf(t)-ktf(t)-k[1-F(t)]+ktf(t)\\ &= (k+c)F(t)-k.\end{aligned}$$

Equating to 0, the minimal expected cost is obtained when you leave t^* minutes before your appointment, where t^* satisfies

$$F(t^*)=\frac{k}{k+c}$$

that is $t^*=F^{-1}(\frac{k}{k+c})$ if F^{-1} exists.
How do we know that this t^* gives us a minimum and not a maximum?

5.3 Uniform Distribution

Definition

A random variable X is said to be **uniformly** distributed over the interval $(0,1)$ if its probability density function is given by

$$f_X(x) = \begin{cases} 1, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}.$$

We denote this by $X \sim U(0,1)$.

$E(X) = \frac{1}{2}$ $\quad$ $Var(X) = \frac{1}{12}$

Finding F_X:

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\,dy = \begin{cases} 0, & \text{if } x < 0 \\ x, & \text{if } 0 \le x < 1 \\ 1, & \text{if } 1 \le x \end{cases}.$$

5.3 Uniform Distribution

Definition

In general, for $a < b$, we say that a random variable X is **uniformly** distributed over the interval (a,b) if its probability density function is given by

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{otherwise} \end{cases}.$$

We denote this by $X \sim U(a,b)$.

In a similar way,

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\,\mathrm{d}y = \begin{cases} 0, & \text{if } x < a \\ (x-a)/(b-a), & \text{if } a \le x < b \\ 1, & \text{if } b \le x \end{cases}.$$

It was shown that

$$E(X) = \frac{a+b}{2} \quad \text{and} \quad \mathrm{var}(X) = \frac{(b-a)^2}{12}.$$

5.3 Uniform Distribution

Example Buses arrive at a specified stop at 15-minute intervals starting at 7 am. That is, they arrive at 7, 7:15, 7:30, 7:45, and so on. If a passenger arrives at the stop at a time that is uniformly distributed between 7 and 7:30, find the probability that he waits

(i) less than 5 minutes for a bus;

(ii) more than 10 minutes for a bus.

Solution

Let X denote the arrival time of the passenger (after 7 in minutes). Then, $X \sim U(0,30)$.

(i) The passenger waits less than 5 minutes for a bus when and only when he arrives (a) between 7:10-7:15 or (b) 7:25-7:30. So the desired probability is

$$P(10 < X < 15) + P(25 < X < 30) = \frac{15-10}{30} + \frac{30-25}{30} = \frac{1}{3}.$$

(ii) The passenger waits more than 10 minutes for a bus when and only when he arrives (a) between 7:00-7:05 or (b) 7:15-7:20. So the desired probability is

$$P(0 < X < 5) + P(15 < X < 20) = \frac{5-0}{30} + \frac{20-15}{30} = \frac{1}{3}.$$

5.4 Normal Distribution

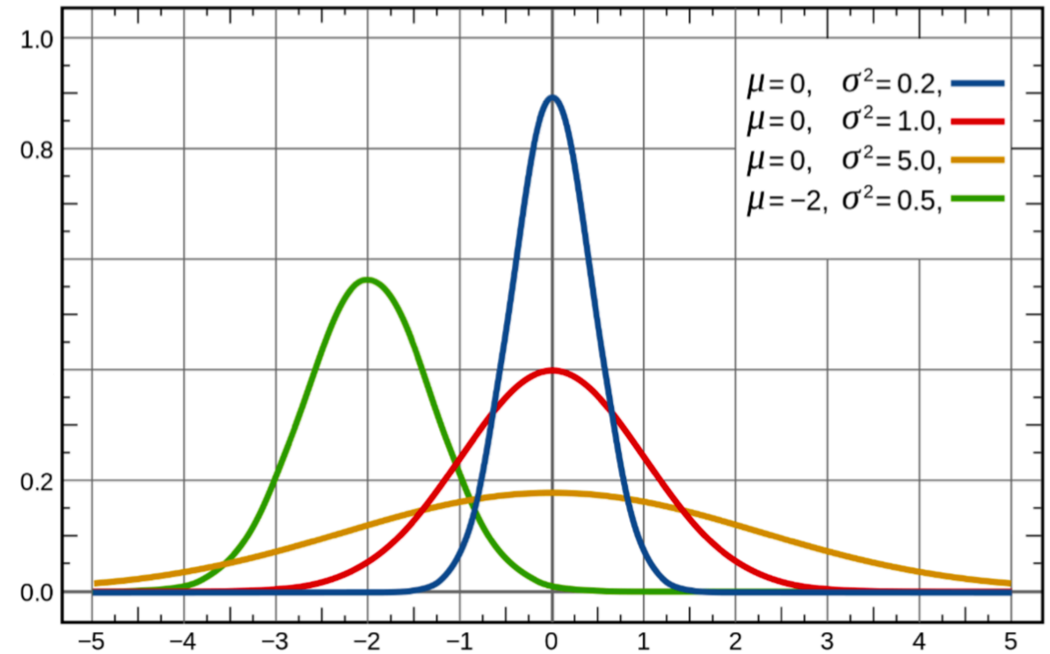
Definition

A random variable X is said to be **normally** distributed with parameters μ and σ^2 if its probability density function is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}, \quad -\infty < x < \infty.$$

We denote this by $X \sim N(\mu, \sigma^2)$.

Note that, this density function is **bell-shaped**, always positive, symmetric at μ and attains its maximum at $x = \mu$.



5.4 Normal Distribution

Let's verify that $f(x)$ is indeed a probability density function, that is,

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}\,\mathrm{dx} = 1.$$

To do so, make the substitution $y = (x-\mu)/\sigma$ to get

$$\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{-(x-\mu)^2/(2\sigma^2)}\,\mathrm{dx} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-y^2/2}\,\mathrm{dy}.$$

Thus we must show that

$$\int_{-\infty}^{\infty} e^{-y^2/2}\,\mathrm{dy} = \sqrt{2\pi}.$$

Towards this end, let $I = \int_{-\infty}^{\infty} e^{-y^2/2}\,\mathrm{dy}$. Then

$$I^2 = \int_{-\infty}^{\infty} e^{-x^2/2}\,\mathrm{dx} \int_{-\infty}^{\infty} e^{-y^2/2}\,\mathrm{dy} \quad = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}\,\mathrm{dy}\,\mathrm{dx}.$$

Perform a change of variable: $x = r\cos\theta$, $y = r\sin\theta$,

$$\begin{aligned} I^2 = \int_0^{\infty}\int_0^{2\pi} e^{-r^2/2} r\,\mathrm{d}\theta\,\mathrm{dr} \quad &= 2\pi \int_0^{\infty} re^{-r^2/2}\,\mathrm{dr} \\ &= 2\pi. \end{aligned}$$

$$\int_0^{\infty} r e^{-\frac{r^2}{2}} dr$$

let $u = -\frac{r^2}{2}$, $du = -r dr$,

$dr = -\frac{1}{r} du$

$$-\int_0^{-\infty} e^u du$$

$$= -[e^u]_0^{-\infty}$$

$$= -(0-1) = 1$$

5.4 Normal Distribution

Definition

A normal random variable is called a **standard normal** random variable when $\mu = 0$ and $\sigma = 1$ and is denoted by Z. That is $Z \sim N(0,1)$. Its probability density function is usually denoted by ϕ and its distribution function by Φ. That is,

$$\begin{aligned}
\phi(x) &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2}; \\
\Phi(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} \, dy .
\end{aligned}$$

Theorem

If $X \sim N(\mu, \sigma^2)$, and then we standardize X, i.e. subtract μ from X and then divide by σ then

$$\frac{X-\mu}{\sigma} \sim N(0,1) .$$

Thus, notationally, we have $Z = \dfrac{X-\mu}{\sigma}$ **.**

5.4 Normal Distribution

Let $Y \sim N(\mu, \sigma^2)$ and $Z \sim N(0,1)$, then

$$\begin{aligned} P(a < Y \leq b) &= P\left(\frac{a-\mu}{\sigma} < Z \leq \frac{b-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right). \end{aligned}$$

This is because

$$\begin{aligned} P(a < Y \leq b) &= \frac{1}{\sqrt{2\pi}\sigma} \int_a^b e^{-\frac{(y-\mu)^2}{2\sigma^2}} \, \mathrm{dy} \\ &= \frac{1}{\sqrt{2\pi}} \int_{(a-\mu)/\sigma}^{(b-\mu)/\sigma} e^{-t^2/2} \, \mathrm{dt} \\ &= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right), \end{aligned}$$

where we let $t = (y-\mu)/\sigma$.

5.4 Normal Distribution

🤔 If $X \sim \mathcal{N}(\mu, \sigma^2)$ for any $\mu, \sigma > 0$

How do we calculate P(a<X<b) for general a, b>0 ?

Don't try to calculate $$\int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}\,dx$$ *by hand*

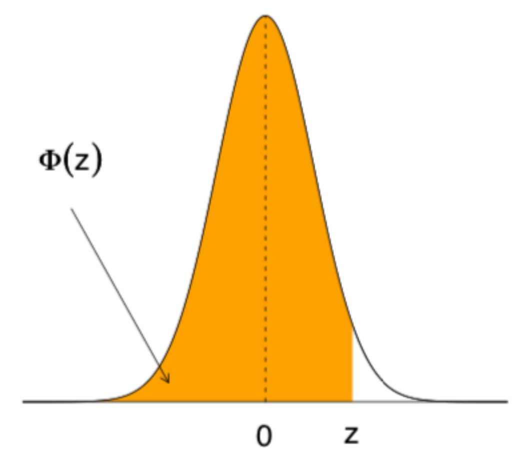
It is simply impossible.

They can be precisely approximated by computer algorithms.

Since any normal r.v. can be standardized, it suffices to provide numerical tables for standard normal r.v.

5.4 Normal Distribution

DISTRIBUTION FUNCTION OF THE NORMAL DISTRIBUTION



The function tabulated is $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}u^2}\,\mathrm{du}.$

	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
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852

$\Phi(0.12) = 0.5478$

$\Phi(0.55) = 0.7088$

5.4 Normal Distribution

(Properties of the Standard Normal)

(i) $P(Z \geq 0) = P(Z \leq 0) = 0.5$.

(ii) $-Z \sim N(0,1)$.

(iii) $P(Z \leq x) = 1 - P(Z > x)$ for $-\infty < x < \infty$.

(iv) $P(Z \leq -x) = P(Z \geq x)$ for $-\infty < x < \infty$.

(v) If $Y \sim N(\mu, \sigma^2)$, then $X = \frac{Y-\mu}{\sigma} \sim N(0,1)$.

(vi) If $X \sim N(0,1)$, then $Y = aX + b \sim N(b, a^2)$ for $a, b \in \mathbb{R}$.

5.4 Normal Distribution

Example When $X \sim N(65, 5^2)$, compute $P(47.5 < X \leq 80)$.

Solution

$$
\begin{aligned}
P(47.5 < X \leq 80) &= P\left(\frac{47.5-65}{5} < Z \leq \frac{80-65}{5}\right) \\
&= P(-3.5 < Z \leq 3) \\
&= P(Z \leq 3) - P(Z \leq -3.5) \\
&= P(Z \leq 3) - P(Z \geq 3.5) \\
&= P(Z \leq 3) - (1 - P(Z < 3.5)) \\
&= 0.99865 - 1 + 0.999767 = 0.998417.
\end{aligned}
$$

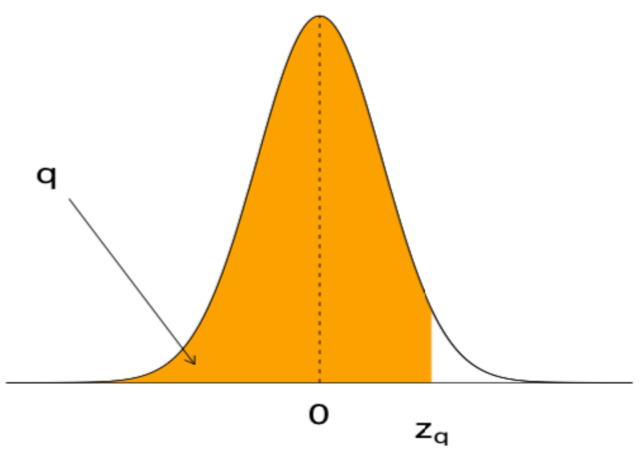
Theorem (a) If $Y \sim N(\mu, \sigma^2)$, then $E(Y) = \mu$ and $\operatorname{var}(Y) = \sigma^2$.

(b) If $Z \sim N(0, 1)$, then $E(Z) = 0$ and $\operatorname{var}(Z) = 1$.

5.4 Normal Distribution

Definition *The qth quantile of a random variable X is defined as a number z_q so that $P(X \leq z_q) = q$.*

Quantiles of the Normal Distribution



For a given q, this table gives z_q such that $\Phi(z_q) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_q} e^{-\frac{1}{2}u^2}\,\mathrm{du} = q$.

q	z_q	q	z_q	q	z_q	q	z_q	q	z_q	q	z_q	q	z_q
0.50	0.000	0.950	1.645	0.960	1.751	0.970	1.881	0.980	2.054	0.990	2.326	0.95	1.645
0.55	0.126	0.951	1.655	0.961	1.762	0.971	1.896	0.981	2.075	0.991	2.366	0.99	2.326
0.60	0.253	0.952	1.665	0.962	1.774	0.972	1.911	0.982	2.097	0.992	2.409	0.999	3.090
0.65	0.385	0.953	1.675	0.963	1.787	0.973	1.927	0.983	2.120	0.993	2.457	0.9999	3.719
0.70	0.524	0.954	1.685	0.964	1.799	0.974	1.943	0.984	2.144	0.994	2.512	0.99999	4.265
0.75	0.674	0.955	1.695	0.965	1.812	0.975	1.960	0.985	2.170	0.995	2.576	0.975	1.960
0.80	0.842	0.956	1.706	0.966	1.825	0.976	1.977	0.986	2.197	0.996	2.652	0.995	2.576
0.85	1.036	0.957	1.717	0.967	1.838	0.977	1.995	0.987	2.226	0.997	2.748	0.9995	3.291
0.90	1.282	0.958	1.728	0.968	1.852	0.978	2.014	0.988	2.257	0.998	2.878	0.99995	3.891
0.95	1.645	0.959	1.739	0.969	1.866	0.979	2.034	0.989	2.290	0.999	3.090	0.999995	4.417

5.4 Normal Distribution

Example The width of a slot of a duralumin in forging is (in inches) normally distributed with $\mu = 0.9000$ and $\sigma = 0.0030$. The specification limits were given as 0.9000 ± 0.0050.

(a) What percentage of forgings will be defective?

(b) What is the maximum allowable value of σ that will permit no more than 1 in 100 defectives when the widths are normally distributed with $\mu = 0.9000$ and σ?

Solution

(a) Let X be the width of our normally distributed slot. The probability that a forging is acceptable is given by

$$P(0.895 < X < 0.905) = P\left(\frac{0.905-0.9}{0.003} < Z < \frac{0.895-0.9}{0.003}\right)$$
$$= P(-1.67 < Z < 1.67) \;= 2\Phi(1.67) - 1 = 0.905.$$

So that the probability that a forging is defective is $1 - 0.905 = 0.095$. Thus 9.5 percent of forgings are defective.

(b) We need to find the value of σ such that $P(0.895 < X < 0.905) = \frac{99}{100}$.

Now $P(0.895 < X < 0.905) = 2P\left(Z < \frac{0.905-0.9}{\sigma}\right) - 1.$

We thus have to solve for σ so that $2P\left(Z < \frac{0.005}{\sigma}\right) - 1 = 0.99.$ or $P\left(Z < \frac{0.005}{\sigma}\right) = (1+0.99)/2 = 0.995.$

The normal quantile table shows that $P(Z < 2.576) = 0.995$ so we can use $\frac{0.005}{\sigma} = 2.576$ which gives $\sigma = 0.0019$.

5.4 Normal Distribution

Example An expert witness in a paternity suit testifies that the length (in days) of pregnancy is approximately normally distributed with parameters $\mu = 270$ and $\sigma = 10$. The defendant in the suit is able to prove that he was out of the country during a period that began 290 days before the birth of the child and ended 240 days before the birth. If the defendant was, in fact, the father of the child, what is the probability that the mother could have had a very long or a very short pregnancy indicated by the testimony?

Solution Let X denote the length of the pregnancy and assume that the defendant is the father. Then the probability of the birth could occur within the indicated duration is

$$\begin{aligned}
& P(X > 290 \text{ or } X < 240) \\
= \;& P(X > 290) + P(X < 240) \\
= \;& P\left(Y > \frac{290-270}{10}\right) + P\left(Y < \frac{240-270}{10}\right) \\
= \;& 1 - \Phi(2) + \Phi(-3) \\
= \;& 1 - \Phi(2) + [1 - \Phi(3)] = 0.0241.
\end{aligned}$$

5.4 Normal Distribution

Example (The 3-σ Rule). Let $X \sim N(\mu, \sigma^2)$. Compute

$$P(|X-\mu| > 3\sigma).$$

Solution

$$\begin{aligned} P(|X-\mu| > 3\sigma) &= P(|Z| > 3) \\ &= 2[1-\Phi(3)] \\ &= 2(1-0.9987) \\ &= 0.0026 = 0.26\%. \end{aligned}$$

This says that for a normal distribution, nearly all (99.74%) of the values lie within 3 standard deviations of the mean.

5.5 Exponential Distribution

Definition

A random variable X is said to be **exponentially** distributed with parameter $\lambda > 0$ if its probability density function is given by

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

The distribution function of X is given by

$$F_X(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-\lambda x}, & x > 0 \end{cases}.$$

Recall the memoryless property of exponential distribution

$$P(X > s+t | X > s) = P(X > t), \quad \text{for } s,t > 0.$$

5.5 Exponential Distribution

Mean and variance of $X \sim \text{Exp}(\lambda)$:

$$E(X) = \frac{1}{\lambda} \quad \text{var}(X) = \frac{1}{\lambda^2}.$$

Example Suppose that the length of a phone call in minutes is an exponential random variable with parameter $\alpha = 1/10$. Someone arrives immediately ahead of you at a public phone booth, find the probability that you will have to wait

(i) more than 10 minutes;

(ii) between 10 to 20 minutes.

Solution

Let X denote the duration of his call.

(i) $$P(X > 10) = 1 - F_X(10) = e^{-10\times 1/10} = e^{-1} = 0.368.$$

(ii) $$P(10 < X < 20) = F(20) - F(10) = (1 - e^{-2}) - (1 - e^{-1}) = 0.233.$$

5.5 Exponential Distribution

??

Example Consider a post office that is staffed by two clerks. Suppose that when Mr. Smith enters the system, he discovers that Ms. Jones is being served by one of the clerks and Mr. Brown by the other. Suppose also that Mr. Smith is told that his service will begin as soon as either Ms. Jones or Mr. Brown leaves. If the amount of time that a clerk spends with a customer is exponentially distributed with parameter λ, what is the probability that, of the three customers, Mr. Smith is the last to leave the post office?

Solution

The answer is obtained by reasoning as follows: Consider the time at which Mr. Smith first finds a free clerk. At this point, either Ms. Jones or Mr. Brown would have just left, and the other one would still be in service. However, because the exponential is memoryless, it follows that the additional amount of time that this other person (either Ms. Jones or Mr. Brown) would still have to spend in the post office is exponentially distributed with parameter λ. That is, it is the same as if service for that person were just starting at this point. Hence, by symmetry, the probability that the remaining person finishes before Smith leaves must equal $\frac{1}{2}$

5.5 Exponential Distribution

Example Jones figures that the total number of thousands of miles that an auto can be driven before it would need to be junked is an exponential random variable with parameter $1/20$. Smith has a used car that he claims has been driven only $10,000$ miles. If Jones purchases the car, what is the probability that she would get at least 20,000 additional miles out of it? Repeat under the assumption that the lifetime mileage of the car is not exponentially distributed but rather is (in thousands of miles) uniformly distributed over $(0, 40)$.

Solution

Let T be the lifetime mileage of the car in thousands of miles. Since the exponential random variable has no memory, the fact that the car has been driven $10,000$ miles makes no difference. The probability we are looking for is

$$P(T > 20) = e^{-\frac{1}{20}(20)} = e^{-1}.$$

If the lifetime distribution is not exponential but is uniform over $(0, 40)$ then the desired probability is given by

$$P(T > 30 | T > 10) = \frac{P(T > 30)}{P(T > 10)} = \frac{(1/4)}{(3/4)} = \frac{1}{3}.$$

5.6 Gamma Distribution

Definition

A random variable X is said to have a **gamma distribution** with parameters (α, λ), denoted by $X \sim \Gamma(\alpha, \lambda)$, if its probability density function is given by

$$f_X(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}, & x \geq 0 \\ 0, & x < 0 \end{cases},$$

where $\lambda > 0$, $\alpha > 0$

$\Gamma(\alpha)$, called the **gamma function**, is defined by

$$\Gamma(\alpha) = \int_0^\infty e^{-y} y^{\alpha-1} \, dy .$$

gamma distribution integration = 1?

$$f(x) = \frac{\lambda e^{-\lambda x} (\lambda x)^{\alpha - 1}}{\Gamma(\alpha)}, \quad x \geq 0$$

$$\int_0^{+\infty} f(x)\,dx = \frac{\int_0^{+\infty} \lambda e^{-\lambda x} (\lambda x)^{\alpha-1}\,dx}{\Gamma(\alpha)} = 1$$

推出

$$\Gamma(\alpha) = \int_0^{+\infty} \lambda e^{-\lambda x} (\lambda x)^{\alpha-1}\,dx$$

但 $\Gamma(\alpha)$ 是

$$\int_0^{+\infty} e^{-y} y^{\alpha-1}\,dy$$

Trick: change of variable: $y = \lambda x$.

$$\Gamma(\alpha) = \int_0^{+\infty} e^{-y} y^{\alpha-1}\,dy.$$

→ cannot evaluate

$$\Gamma(\alpha) = \int_0^{+\infty} e^{-y} y^{\alpha-1} dy$$

$$\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$$

Proof of above statement:

$$\Gamma(\alpha-1) = \int_0^{+\infty} e^{-y} y^{\alpha-2} dy.$$

let $u = \alpha - 2$, $du = d\alpha$

$J0$

Integration of $\Gamma(\alpha)$ by parts yields

$$\begin{aligned}
\Gamma(\alpha) &= -e^{-y}y^{\alpha-1}\Big|_0^{\infty} + \int_0^{\infty} e^{-y}(\alpha - 1)y^{\alpha-2}\,dy \\
&= (\alpha - 1)\int_0^{\infty} e^{-y}y^{\alpha-2}\,dy \qquad (6.1) \\
&= (\alpha - 1)\Gamma(\alpha - 1)
\end{aligned}$$

5.6 Gamma Distribution

Remarks

$$f_X(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}, & x \geq 0 \\ 0, & x < 0 \end{cases},$$

(a) $\Gamma(1) = \int_0^\infty e^{-y}\,\mathrm{dy} = 1.$

proof of this?

(b) It can be shown, via integration by parts, that $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$.

(c) For integral values of α, say, $\alpha = n$,

$$\begin{aligned} \Gamma(n) &= (n-1)\Gamma(n-1) \\ &= (n-1)(n-2)\Gamma(n-2) \\ &= \ldots \quad = (n-1)(n-2)\cdots 3\cdot 2\cdot \Gamma(1) \quad = (n-1)! \end{aligned}$$

(d) Note that $\Gamma(1,\lambda) = \mathrm{Exp}(\lambda)$.

(e) $\Gamma\left(\frac{1}{2}\right) = \int_0^\infty e^{-y} y^{-\frac{1}{2}}\,\mathrm{dy} = \sqrt{\pi}.$

5.6 Gamma Distribution

Gamma distribution arises naturally in processes for which the waiting times between events are relevant.
It can be thought of as a waiting time between Poisson distributed events.

The gamma distribution can be used a range of disciplines including queuing models, climatology, and financial services.

Examples of events that may be modeled by gamma distribution include:

- The amount of rainfall accumulated in a reservoir
- The size of loan defaults or aggregate insurance claims
- The flow of items through manufacturing and distribution processes
- The load on web servers
- The many and varied forms of telecom exchange

5.7 Beta Distribution

Definition

A random variable X is said to have a **beta distribution** with parameters (a, b), denoted by $X \sim \text{Beta}(a, b)$, if its density is given by

$$f(x) = \begin{cases} \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

where

$$B(a, b) = \int_0^1 x^{a-1}(1-x)^{b-1}\,\mathrm{dx}$$

is known as the **beta function**.

It can be shown that

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

Using this, it is easy to show that if X is $\text{Beta}(a, b)$, then

$$E[X] = \frac{a}{a+b} \quad \text{and} \quad \text{var}(X) = \frac{ab}{(a+b)^2(a+b+1)}.$$

5.8 Cauchy Distribution

Definition

A random variable X is said to have a **Cauchy distribution** with parameter θ, $-\infty < \theta < \infty$, denoted by $X \sim \text{Cauchy}(\theta)$, if its density is given by

$$f(x) = \frac{1}{\pi} \frac{1}{1 + (x - \theta)^2}, \quad -\infty < x < \infty.$$

$$E(X) = \infty, \quad \text{Var}(X) = +\infty$$

5.9 Approximations of Binomial Random Variables

Let $X \sim \text{Bin}(n, p)$. Here n is assumed (which is natural) to be large (say, ≥ 30).
There are two commonly used approximations of the binomial distributions:

(a) Normal approximation; and

(b) Poisson approximation.

5.9 Approximations of Binomial Random Variables

Normal approximation of Bin(n, p)

Theorem (De Moivre-Laplace Limit Theorem). *Suppose that $X \sim \text{Bin}(n, p)$. Then for any $a < b$,*

$$P\left(a < \frac{X - np}{\sqrt{npq}} \leq b\right) \to \Phi(b) - \Phi(a)$$

That is,

$$\text{Bin}(n, p) \approx N(np, npq).$$

Equivalently,

$$\frac{X - np}{\sqrt{npq}} \approx Z$$

where $Z \sim N(0, 1)$.

Remarks The normal approximation will be generally quite good for values of n satisfying $np(1-p) \geq 10$.

5.9 Approximations of Binomial Random Variables

Normal approximation of Bin(n, p)

Approximation is further improved if we incorporate

(Continuity-correction) If $X \sim \text{Bin}(n, p)$, then

$$P(X = k) = P\left(k - \frac{1}{2} < X < k + \frac{1}{2}\right)$$

$$P(X \geq k) = P\left(X \geq k - \frac{1}{2}\right)$$

$$P(X \leq k) = P\left(X \leq k + \frac{1}{2}\right)$$

5.9 Approximations of Binomial Random Variables

Normal approximation of Bin(n, p)

Example Let X be the number of times that a fair coin, flipped 40 times, lands heads. Find the probability that $X = 20$. Use the normal approximation to compute this probability as well and compare the two answers.

Solution

Exact answer:

$$P(X = 20) = \binom{40}{20}\left(\frac{1}{2}\right)^{40} = 0.1254.$$

Approximate answer:

$$\begin{aligned} P(X = 20) &= P(19.5 \le X \le 20.5) \\ &= P\left(\frac{19.5 - 20}{\sqrt{10}} \le \frac{X - 20}{\sqrt{10}} \le \frac{20.5 - 20}{\sqrt{10}}\right) \\ &\approx P(-0.16 \le Z \le 0.16) \\ &= 0.1272. \end{aligned}$$

5.9 Approximations of Binomial Random Variables

Normal approximation of Bin(n, p)

Example Let $X \sim \text{Bin}(1000000, 0.01)$. We are interested in

$$P(8000 \leq X \leq 120000).$$

Solution

5.9 Approximations of Binomial Random Variables

Normal approximation of Bin(n, p)

An example on how to apply Continuity Correction:
Let $X \sim \text{Bin}(60, 0.3)$ and $Z \sim N(0,1)$. Then

(i) $P(12 \leq X \leq 26) \approx P\left(\frac{11.5-18}{\sqrt{12.6}} \leq Z \leq \frac{26.5-18}{\sqrt{12.6}}\right)$.

(ii) $P(12 < X \leq 26) \approx P\left(\frac{12.5-18}{\sqrt{12.6}} \leq Z \leq \frac{26.5-18}{\sqrt{12.6}}\right)$.

(iii) $P(12 \leq X < 26) \approx P\left(\frac{11.5-18}{\sqrt{12.6}} \leq Z \leq \frac{25.5-18}{\sqrt{12.6}}\right)$.

(iv) $P(12 < X < 26) \approx P\left(\frac{12.5-18}{\sqrt{12.6}} \leq Z \leq \frac{25.5-18}{\sqrt{12.6}}\right)$.

5.9 Approximations of Binomial Random Variables

Poisson approximation of Bin(n, p)

The Poisson distribution is used as an approximation to the binomial distribution when the parameters n and p are large and small, respectively and that np is moderate.

As a working rule, use the Poisson approximation if $p < 0.1$ and put $\lambda = np$.

Example Let $X \sim \text{Bin}(400, 0.01)$, then $X \approx Z$ where $Z \sim \text{Poisson}(\lambda)$ and $\lambda = np = 4$. Therefore,

$$0.890375 = P(X \leq 6) \approx P(Z \leq 6) = 0.889326.$$

5.10 Distribution of a Function of a Random Variable

Example Let $X \sim N(0,1)$. What are the distribution function and probability density function of Y, where $Y = X^2$?

Solution

First note that Y takes nonnegative values. Therefore for $y > 0$, we have

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(X^2 \leq y) \\ &= P(-\sqrt{y} \leq X \leq \sqrt{y}) = \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{y}} e^{-x^2/2} \mathrm{d}x\,. \end{aligned}$$

Hence $f_Y(y) = F_Y'(y) = \frac{2}{\sqrt{2\pi}} \frac{1}{2\sqrt{y}} e^{-y/2} = \frac{y^{-1/2} e^{-y/2}}{\sqrt{2\pi}}.$

This gives $F_Y(y) = \begin{cases} 0, & y \leq 0 \\ \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{y}} e^{-x^2/2} \mathrm{d}x, & y > 0 \end{cases}$

and $f_Y(y) = \begin{cases} 0, & y \leq 0 \\ \frac{y^{-1/2} e^{-y/2}}{\sqrt{2\pi}}, & y > 0 \end{cases}.$

Y is known in the literature as a **chi-square (χ^2) random variable** of degree 1, denoted by χ_1^2. Note that χ_1^2 is $\Gamma\left(\frac{1}{2}, \frac{1}{2}\right)$.

5.10 Distribution of a Function of a Random Variable

Example Let $X \sim N(0,1)$. Define $Y = e^X$, commonly known as the **lognormal random variable**. Find the probability density function f_Y .

Solution

First note that Y takes nonnegative values. Therefore for $y > 0$, we have

$$\begin{aligned} F_Y(y) &= P(Y \le y) = P(e^X \le y) \\ &= P(X \le \ln y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\ln y} e^{-x^2/2} \mathrm{dx}. \end{aligned}$$

Hence

$$f_Y(y) = F_Y'(y) = \frac{1}{\sqrt{2\pi}} \frac{1}{y} e^{-(\ln y)^2/2}.$$

This gives

$$f_Y(y) = \begin{cases} 0, & y \le 0 \\ \frac{1}{y\sqrt{2\pi}} e^{-(\ln y)^2/2}, & y > 0 \end{cases}.$$

5.10 Distribution of a Function of a Random Variable

Example Let X be a continuous random variable with probability density function f_X. And let $Y = X^n$ where n is odd. Find the probability density function f_Y.

Solution

Let $y \in \mathbb{R}$, then

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(X^n \leq y) \\ &= P(X \leq y^{1/n}) = F_X(y^{1/n}). \end{aligned}$$

Hence

$$f_Y(y) = F_Y'(y) = \frac{1}{n} y^{1/n-1} f_X(y^{1/n}).$$

5.10 Distribution of a Function of a Random Variable

Theorem *Let X be a continuous random variable having probability density function f_X. Suppose that $g(x)$ is a strictly monotonic (increasing or decreasing), differentiable (and thus continuous) function of x. Then the random variable Y defined by $Y = g(X)$ has a probability density function given by*

$$f_Y(y) = \begin{cases} f_X\left(g^{-1}(y)\right)\left|\frac{d}{dy}g^{-1}(y)\right|, & \text{if } y = g(x) \text{ for some } x \\ 0, & \text{if } y \neq g(x) \text{ for all } x \end{cases}$$

where $g^{-1}(y)$ is defined to be equal that value of x such that $g(x) = y$.

Proof We shall assume that $g(x)$ is an increasing function. Suppose $y = g(x)$ for some x. Then, with $Y = g(X)$,

$$F_Y(y) = P(g(X) \leq y) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y)).$$

Differentiation gives

$$f_Y(y) = f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y),$$

which agrees with the form given in the theorem, since $g^{-1}(y)$ is nondecreasing, so its derivative is non-negative.
When $y \neq g(x)$ for any x, $F_Y(y)$ is either 0 or 1. In either case $f_Y(y) = 0$. $\square$

X, $f(x)$, $Y = g(x)$

① compute the cdf $F_Y(y) = P(Y \le y) = P(g(x) \le y)$

$= P(X \le g^{-1}(y))$

5.10 Distribution of a Function of a Random Variable

Example Let Y be a continuous random variable with distribution function F. We assume further that F is a strictly increasing function. Define the random variable X by $X = F(Y)$.

(i) What are the possible values of X?

(ii) Find the probability density function of X. Can you identify X?

Solution

(i) X takes values in $[0,1]$. Since F is cdf, take value in $[0,1]$.

(ii) For $x \in (0,1)$,

$$\begin{aligned} F_X(x) &= P(X \le x) \\ &= P(F(Y) \le x) \\ &= P(Y \le F^{-1}(x)) \quad \leftarrow \text{true only } F(Y) \text{ is strictly increasing.} \\ &= F(F^{-1}(x)) \\ &= x. \end{aligned}$$

$$f_X(x) = \begin{cases} 1 & \text{if } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases},$$

that is, $X \sim \text{uniform}(0,1)$.

5.10 Distribution of a Function of a Random Variable

Remark There are two things worth taking note of in *previous example*

(a) If X is a continuous random variable with distribution function F, then $F(X)$ is the uniform distribution $U(0,1)$.

(b) If U is the uniform distribution $U(0,1)$, then $F^{-1}(U)$ will have the distribution function $F(x)$. This result, known as the *inverse transformation method*, is often made used of to generate continuous random variables having distribution function F in computer packages.

5.10 Distribution of a Function of a Random Variable

Example (Generating an Exponential Random Variable). Let

$$F(x) = 1 - e^{-\lambda x}, \quad x \geq 0$$

be the distribution function of an exponential random variable with parameter λ. Then $F^{-1}(u)$ is that value x such that

$$u = F(x) = 1 - e^{-\lambda x}$$

or, equivalently,

$$x = -\frac{1}{\lambda}\log(1-u)$$

So we can generate an exponential random variable X with parameter λ by generating a uniform $(0,1)$ random variable U and setting

$$X = -\frac{1}{\lambda}\log(1-U).$$

Because $1-U$ is also a uniform $(0,1)$ random variable, it follows that $-\frac{1}{\lambda}\log(1-U)$ and $-\frac{1}{\lambda}\log(U)$ have the same distribution, thus showing that

$$X = -\frac{1}{\lambda}\log(U)$$

is also exponential with parameter λ.