MH2500 Probability and Introduction to Statistics

Handout 2 - Random Variables

Synopsis

After the introduction to probability in Chapter 1, we move on to introduce discrete and continuous random variables, and the various distributions. The discussions on continuous random variables require knowledge in calculus.

- Discrete Random Variables, probability mass function, cumulative distribution function (cdf),
  - Bernoulli Random Variables,
  - Binomial distribution,
  - hypergeometric distribution,
  - poisson distribution,
- Continuous Random Variables,
  - Exponential Density
  - Gamma Density
  - Normal Distribution,
- Functions of a Random Variable,

Discrete Random Variables

A **random variable** is just

E.g. Two dice are rolled, and the sample space is

$$\Omega = \{(1,1),(1,2),\ldots,(1,6),(2,1),\ldots,(2,6),\ldots,(6,1),\ldots,(6,6)\}.$$

Suppose we count (i) number on the first die, (ii) sum of the two dice, (iii) product of the two dice. Since the outcome in Ω is random, each of these numbers are random variables.

More formally, a **random variable** is a function

$$X : \Omega \to \mathbb{R}.$$

A **discrete random variable** is a random variable that can only take a finite or at most countably infinite number of values.

Probability Mass Function

Definition

Given a discrete random variable X taking values $x_1, x_2, \ldots$, its **probability mass function** or the **frequency function** of a random variable X is the function p where

$$p(x_i) = P(X = x_i) \qquad \text{and} \qquad \sum_{i=1}^{\infty} p(x_i) = 1.$$

. . . .

Example: A coin is flipped three times. Then the sample space is

$$\Omega = \{hhh, hht, hth, htt, thh, tht, tth, ttt\},$$

where h denotes head and t denotes tail. Let X be the total number of heads.

Example continue

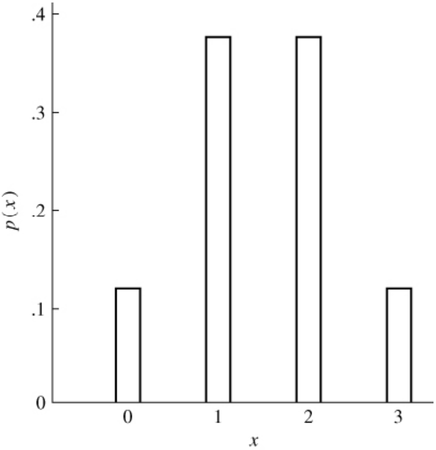


Figure 2.1 **A probability mass function.**

Cumulative distribution function (cdf)

Definition

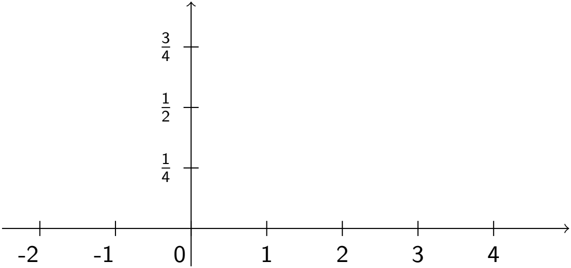
The **cumulative distribution function (cdf)** of a random variable is defined as the function $F : \mathbb{R} \to [0, 1]$ where

$$F(x) = P(X \leq x), \qquad -\infty < x < \infty.$$

Note: we use uppercase for cumulative distribution function and lowercase letters for probability mass function.

CDF

Example. Graph the cdf of the previous example.



Note that we always have

$$\lim_{x\to-\infty} F(x) = 0 \qquad \text{and} \qquad \lim_{x\to\infty} F(x) = 1.$$

Independent

Definition

Suppose X and Y are two discrete random variable taking possible values $x_1, x_2, \ldots$, and $y_1, y_2, \ldots$, respectively. We say that X and Y are **independent** if for all i and j,

$$P(X = x_i \text{ and } Y = y_j) = P(X = x_i)P(Y = y_j).$$

This definition extends to the case where there are more than two variables. For example, we say that the discrete random variables X, Y, and Z are mutually independent if for all i, j, and k,

$$P(X = x_i,\ Y = y_j \text{ and } Z = z_k) = P(X = x_i)P(Y = y_j)P(Z = z_k).$$

Example

A fair coin is flipped three times. Let X be the number of heads in the first two flips, Y be the number of tails in the third flip, and Z be the total number of heads in the three flips.
Prove that (i) X and Y are independent, (ii) X and Z are not independent.

Solution: (i) Since the outcome of the first two coin flips does not affect the outcome of the third flip, it is obvious that X and Y are independent.
But to be rigorous, we compute the probability of
$P(X = x_i \text{ and } Y = y_j) = P(X = x_i)P(Y = y_j)$ for all i and j.

$Y \diagdown X$	0	1	2
0			
1			

(ii)

Bernoulli Random Variable

Definition

A Bernoulli random variable takes on only two values: 1 and 0, with probabilities p and $1-p$. In other words,

$$p(1) = p, \qquad p(0) = 1-p, \qquad p(x) = 0 \quad \text{if } x \neq 0, 1.$$

An alternative representation is

$$p(x) = \begin{cases} p^x(1-p)^{1-x}, & \text{if } x = 0 \text{ or } x = 1; \\ 0, & \text{otherwise.} \end{cases}$$

If A is an event, then the **indicator random variable**, I_A takes the value 1 if A occurs and 0 if A does not occur, i.e.,

$$I_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A; \\ 0, & \text{if } \omega \notin A. \end{cases}$$

Binomial Distribution

Suppose n independent experiments or trials are performed, where n is fixed, and each experiment results in either "success" with probability p or "failure" with probability $1 - p$.

Let X be the total number of successes. Then X is a **binomial random variable** with parameters n and p.

The probability of having k successes is

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

. . . .

Example: Roll a die 5 times and let X count the number of times 6 shows up. Then

$$P(X = 3) = \qquad\qquad\qquad\qquad .$$

Example

Let X be a binomial random variable with $n = 3$ and $p = 0.2$. Find the frequency function of X. For what value of X is the probability the highest?

Solution:

$p(X = 0) =$

$P(X = 1) =$

$P(X = 2) =$

$P(X = 3) =$

The probability is highest at $X =$

Geometric Distribution

The **geometric distribution** is also constructed from independent Bernoulli trials, but from an infinite sequence. On each trial, a success occurs with probability p.

Let X be the total number of trials up to and including the first success. If $X = k$, then there must be $k - 1$ failures preceding the first success.

Hence the probability is

$$P(X = k) = (1 - p)^{k-1} p, \qquad k = 1, 2, \ldots .$$

Note that the sum of probabilities is 1 because

$$\sum_{k=1}^{\infty} P(X = k) \ = \ p \sum_{j=0}^{\infty} (1 - p)^j = \qquad\qquad\qquad .$$

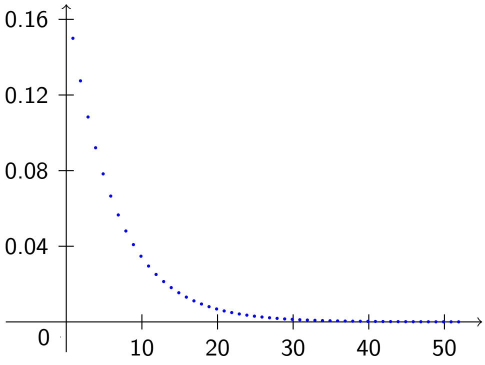
Example

A student approaches people on the street for donations on a flag day. For each person that he approaches randomly, there is a 15% chance that the person donates. What is the probability that the first person who donated was the 10th person he approached?

Solution:
Let X be the number of people the student approached before receiving the first donation.

$$P(X = 10) = 0.85^9 \times 0.15 = 0.0347.$$

Example



The probability mass function of a geometric random variable with $p = 0.85$.

. . . .

Negative Geometric Distribution

The **negative binomial distribution** arises as a generalization of the geometric distribution.

Suppose that a sequence of independent trials, each with probability of success p, is performed until there are r success in all. Let X denote the total number of trials.

If $X = k$, it implies that there are $r - 1$ success in the first $k - 1$ trials and the k-th trial is also a success. This gives

$$P(X = k) = \qquad\qquad\qquad\qquad .$$

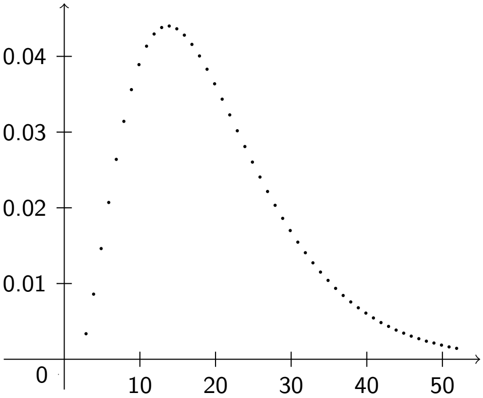
The binomial distribution is distribution of $\#$ successes in fixed n trials.

The negative binomial distribution is the distribution of $\#$ trials needed to obtain a (fixed) r successes.

Example

A student approaches people on the street for donations on a flag day. For each person that he approaches randomly, there is a 15% chance that the person donates. What is the probability that the third person who donated was the 10th person he approached?

Example



The probability mass function of a negative binomial random variable with $p = 0.85$ and $r = 3$.

Hypergeometric Distribution

Suppose an urn contains n balls, of which r are black and $n-r$ are white. Let X denote the number of black balls drawn when taking m balls without replacement. This was discussed in Handout 1, and we know that

$$P(X=k)=\frac{\binom{r}{k}\binom{n-r}{m-k}}{\binom{n}{m}}.$$

Then X is a **hypergeometric random variable** with parameters r, n, and m, and X has a hypergeometric distribution.

Example

In a certain lottery, a player picks six out of 48 numbers and the lottery officials later choose 6 numbers at random. Let X be the number of matches. Find the frequency function of X.

Solution:

$$P(X = k) = \qquad .$$

$x_i =$	0	1	2	3	4	5	6
$P(X = x_i)$	0.427	0.416	0.137	0.0187	0.00105	2.05×10^{-5}	8.15×10^{-8}

Poisson Distribution

The **Poisson frequency function** with parameter λ ($\lambda > 0$) is

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

The frequency function sums to 1 since

$$\sum_{k=0}^{\infty} P(X = k) =$$

Properties of Poisson

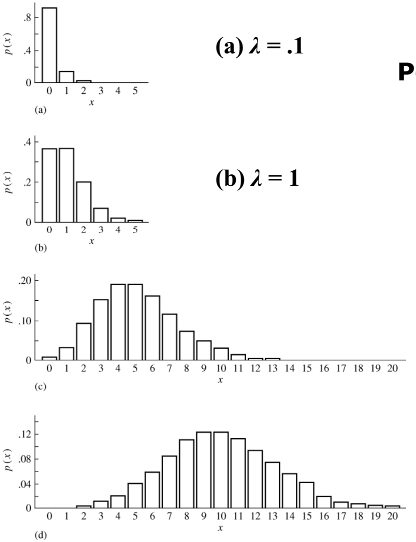


Figure 2.6
Poisson frequency functions.

See also
`http://demonstrations.wolfram.com/PoissonDistribution/`

Poisson Distribution

The Poisson distribution can be derived as the limit of a binomial distribution with $np = \lambda$, as $n \to \infty$.

Proof: For a fixed n, we have $p = \frac{\lambda}{n}$ and so

$$
\begin{aligned}
p(k) &= \\
&= \frac{n!}{k!(n-k)!}\left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \qquad\qquad\qquad\qquad .
\end{aligned}
$$

Proof con't

As $n \to \infty$,

$$\frac{n!}{(n-k)!}\frac{1}{n^k} \to$$

$$\left(1 - \frac{\lambda}{n}\right)^n \to$$

$$\left(1 - \frac{\lambda}{n}\right)^{-k} \to \quad .$$

Thus $p(k) \to \lambda^k e^{-\lambda}/k!$.

Example C

Suppose That an office receives telephone calls as a Poisson process with $\lambda = 0.5$ per min. Find
(i) The probability that there are no calls in a 5 min interval
(ii) The probability of exactly one call in a 5 min interval.

Solution:
The number of calls in a 5 min interval follows a Poisson distribution with parameter $\omega = 5\lambda = 2.5$.

(i) The probability of no calls in a 5 min interval is

(ii) The probability of one call in a 5 min interval is

Continuous Random Variables

Recall that formally, a random variable is a function

$$X : \Omega \to \mathbb{R}.$$

and a discrete random variable is a random variable that can only take
number of values.

If Ω is a "continuous" set, then X takes continuous values and so is **continuous random variable**.

For a continuous random variable, the role of the frequency function is taken by a **density function**, $f(x)$, which has the property that $f(x) \geq 0$.

Then f is piecewise continuous and $\int_{-\infty}^{\infty} f(x)dx = 1$.

If X has density function f, then probability that $X \in (a, b)$ is

$$P(a < X < b) = \int_a^b f(x)dx \qquad \text{(area under } f\text{)}.$$

Example A

A **uniform random variable** on the interval [0,1] is a model for what we mean when we say "choose a number at random between 0 and 1."

The density function in this case, a.k.a. the **uniform density** on $[0,1]$, is

$$f(x) = \begin{cases} 1, & \text{if } 0 \le x \le 1 \,; \\ 0, & \text{if } x < 0 \text{ or } x > 1. \end{cases}$$

In general, the uniform density on $[a,b]$ is

$$f(x) = \begin{cases} \quad\quad , & \text{if } a \le x \le b; \\ 0, & \text{if } x < a \text{ or } x > b, \end{cases}$$

and the cumulative distribution function is

$$F(x) = \int_{-\infty}^{x} f(u)du =$$

Remarks

- For a continuous random variable X and $c \in \mathbb{R}$,

$$P(X = c) = \int_c^c f(x)dx = 0.$$

- For the same reason,

$$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b).$$

- For a small $\delta > 0$, if f is continuous at x, then

$$P\left(x = \frac{\delta}{2} \leq X \leq x + \frac{\delta}{2}\right) \; = \; \int_{x=\delta/2}^{x+\delta/2} f(u)du \;\approx\; \delta f(x).$$

("think" of area of rectangle)

Remarks

- The cdf of X is

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(u)du.$$

From this, we have

$$P(a \le X \le b) = \int_{a}^{b} f(u)du = F(b) - F(a).$$

Example

Give the cdf of a uniform random variable on [1,4].

Solution:

$$F(x) = \begin{cases} \quad , & \text{if } x < 1; \\ \quad , & \text{if } 1 \le x \le 4; \\ \quad , & \text{if } x > 4. \end{cases}$$

Inverse of cdf

Suppose X is a continuous random variable with cdf F strictly increasing on some interval $I = [a, b]$ with $F(x) = 0$ for $x \leq a$ and $F(x) = 1$ for $x \geq b$.

Then the inverse function F^{-1} : is well defined.

For any $p \in [0, 1]$, the p-**th quantile** of the distribution F is define to be the value x_p such that $F(x_p) = p$ or $P(X \leq x_p) = p$.
(Under the assumption above, F^{-1} exists and so $x_p = F^{-1}(p)$ is uniquely defined.)

When $p = 1/2$, the $1/2$-quantile is known as the **median** of F. The $1/4$-quantile and $3/4$-quantile are known as the lower and upper **quartiles** or the first and third quartile of F.

Example C

Suppose that $F(x) = x^2$ for $0 \le x \le 1$, i.e.,

$$F(x) = \begin{cases} \quad , & \text{if } x < 0; \\ \quad , & \text{if } 0 \le x \le 1; \\ \quad , & \text{if } x > 1. \end{cases}$$

Find (i) F^{-1}, (ii) $F^{-1}(.5)$, (iii) the median, (iv) the lower quartile, (v) the upper quartile.

Solution:

(i)

(ii) (iii) (iv) (v)

Exponential Density

The exponential density function is

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

The cumulative distribution function is

$$F(x) = \int_{-\infty}^{x} f(u)du = \begin{cases} 1 - e^{-\lambda x}, & \text{if } x \geq 0; \\ 0, & \text{if } x < 0. \end{cases}$$

Like the Poisson distribution, the exponential density depends on a single parameter $\lambda > 0$.

Properties of the Exponential distribution

- Related to Poisson.
  - It is used to model between the occurrence of events in an interval of time.
  - E.g. The time between phone calls received at a help center,

    The time between failures of a system, etc,

    In general, “the time between two consecutive events where the occurrence of events follow a Poisson distribution”.
- Memoryless.
  - The memoryless character of the exponential distribution follows from its relation to a Poisson process.

Exponential: Relation with Poisson

Suppose events occur in unit time as a Poisson process with parameter λ and that an event occurs at time t_0.

(Recall: If X is the number of events in unit time, $P(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}$, and so $P(X = 0) = e^{-\lambda}$.)

Instead studying the probability of the number of events in a certain time, we consider T the length of time until the next event. Then

$$P(T > t) = P(\text{no events in } (t_0, t_0 + t)) = P(Y = 0),$$

where Y is the number of events during time $(t_0, t_0 + t)$, (i.e., in time t).

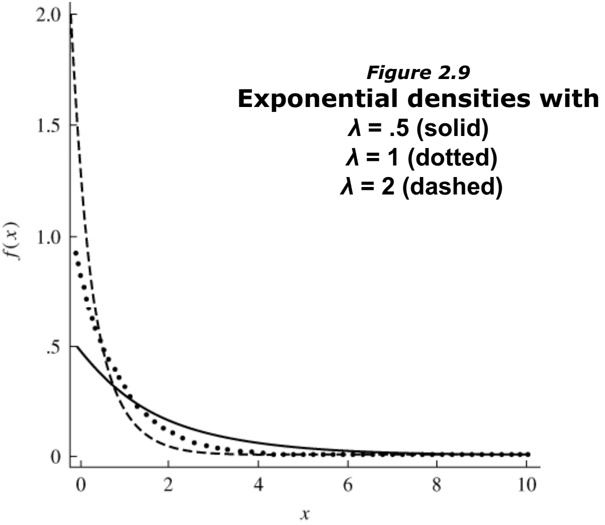
Then Y is a Poisson process with parameter λt. Hence

$$P(T > t) = P(Y = 0) = e^{-\lambda t}.$$

I.e., $\qquad F(t) = \qquad\qquad\qquad .$

Comparing different values of λ

Note that as λ increases, the density drops off more rapidly.



Suppose T is the lifetime of an electronic component. We model T using an exponential distribution with parameter λ.
(i) Find the median (the cdf of T).
(ii) Suppose the electronic component has lasted a length of time s. Find the probability that it will last at least another length of time t.

Solution:
(i) Let η be the median. Then $\qquad = \frac{1}{2}$, which implies

$$\eta = \qquad .$$

(ii)

$$P(T \geq t + s | T \geq s) = \frac{P(T > t + s \text{ and } T > s)}{P(T > s)}$$

$$=$$

$$= \qquad .$$

Exponential: Memoryless

The answer in (ii) does not depend on s. Exponential distribution is **memoryless**.

The exponential distribution is not a good model for human lifetime. E.g. the probability that a 20 year old will live at least another 10 years is not the same as the probability that an 80 year old will live at least another 10 years.

It can be shown that a memoryless distribution follows an exponential distribution.

Gamma Density

The **gamma function** is defined as

$$\Gamma(x) = \int_0^\infty u^{x-1} e^{-u} du, \qquad x > 0.$$

The **gamma density function** depends on two parameters, $\alpha > 0$ and $\lambda > 0$ and is defined as

$$g(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}, \qquad t > 0,$$

and g(t)=0 if $t \leq 0$.

Check that

$$\int_0^\infty g(t) dt = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-\lambda t} dt = 1$$

Properties of the Gamma distribution

Two parameters, α and λ

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

  and the Gamma distribution reduces to

- Recall, suppose $X \sim Po(\lambda)$ and Y counts the number of events from time t_0, then the time T_1 between t_0 to the first event follows an exponential distribution.
  The time T_n between t_0 to the n-th event follows a

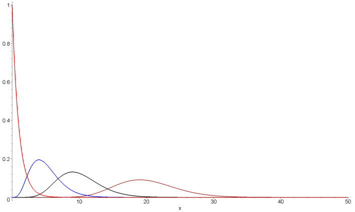
  .

- Applications: used to model arrival times in the Poisson process, e.g. time for failures in systems, physical quantities taking positive values such as rainfalls, and in finances such as size of insurance claims.

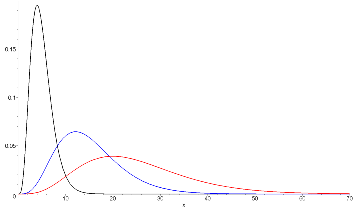
- It is also a generalization of the χ^2 distribution. (Later)

Gamma Distribution: Parameters

Perimeter α, is called the **shape parameter**. Varying α changes the shape of the pdf.
Gamma densities with $\lambda = 1$ and $\alpha = 1, 5, 10, 20$ (red, blue, black, brown)



Parameter λ is called a **rate parameter**, varying which changes the height and spread but not the shape.
Gamma densities with $\alpha = 5$ and $\lambda = 0.2, 0.33, 1$ (red, blue, black)



Note that in some books, gamma density is defined with λ replaced with $1/\lambda$, then λ is known as a **scale parameter**.

Example

Observed times separating a sequence of small earthquakes. Which is a better fit? (Recall that exponential model is memoryless.)

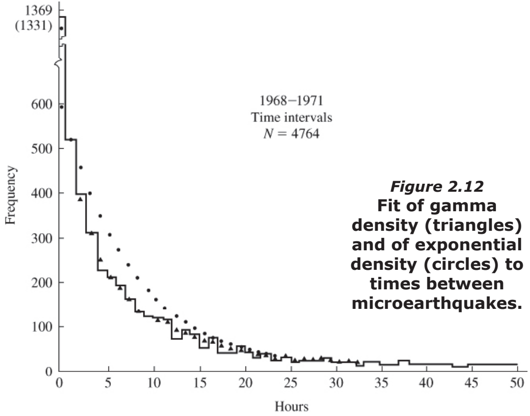


Figure 2.12
Fit of gamma density (triangles) and of exponential density (circles) to times between microearthquakes.

Beta Density

The **beta density** has two parameters a and b and is defined by

$$f(u) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^{a-1}(1-u)^{b-1}, \qquad 0 \leq u \leq 1.$$

It is useful for modelling random variables that are restricted to the interval [0,1].

Beta Density

Suppose we are given two independent random variables X_1 and X_2, where X_1 has a gamma distribution with parameters α and λ while X_2 has gamma distribution with β and λ.

We construct a new random variable

$$X = \frac{X_1}{X_1 + X_2}.$$

Then X measures the proportion of $X_1 + X_2$ that is X_1.

It can be shown that X has the **beta density** with parameters α and β.

Application: Model the behaviour of random variables limited to intervals of finite length.

Normal Distribution

The density function of the **normal distribution** depends on two parameters, μ and σ (where $-\infty < \mu < \infty, \sigma > 0$), is defined by

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

The parameters μ and σ are called the **mean** and **standard deviation** of the normal density.

- Plays central role in probability and statistics. (Will be apparent in later chapters.)
- Also called the Gaussian distribution.
- In Chap 6, we discuss the central limit theorem (CTL), which justifies using the normal distribution in many applications.
  CTL: the sum of a large number of independent random variables is approximately normally distributed.

Normal Distribution

- The cdf cannot be evaluated in closed form and have to be found numerically. (For this course, we just look up tables.)

- However, we can show that (see Ex. 2.5.51)

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

- A shorthand notation for "X follows a normal distribution with parameters μ and σ" is

$$X \sim N(\mu, \sigma^2).$$

- For $\mu = 0$ and $\sigma = 1$, we have the **standard normal density**, $N(0, 1)$.

Normal Distribution

The density function is symmetric about μ, and so

$$f(\mu - x) = f(\mu + x).$$

The standard deviation, σ, determines the rate at which the density falls off.

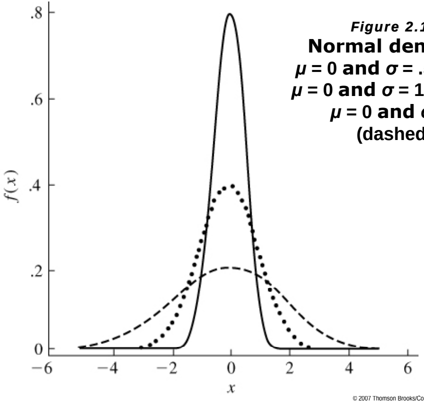


Figure 2.13
Normal densities
$\mu = 0$ and $\sigma = .5$ (solid)
$\mu = 0$ and $\sigma = 1$ (dotted)
$\mu = 0$ and $\sigma = 2$ (dashed)

Examples

Examples where normal distribution is a good model. (Refer to the book for details.)

Example A: Amplitude of background noise in the ocean.

Example B: The velocity of turbulent air flow at a point.

Example C: The average return for investments per day.

Functions of a Random Variable

Suppose X is a random variable with density function $f(x)$ and Y is a random variable that can be expressed in terms of X, say, $Y = g(X)$ for some function g. We would like to find the density function of Y.

E.g.1. X is the number of heads in flipping a coin 10 times, and Y is the number of tails.

E.g.2. X is amount of electricity used by a household and Y is the amount paid.

We illustrate the technique to find the density function for Y.

Suppose $Y = g(X)$, let f_X, F_X denote the density function and cdf of X, respectively, and let f_Y, F_Y be defined likewise. If g^{-1} exists, then

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

Differentiate $F_X(g^{-1}(y))$ with respect to y (via chain rule) to get $f_Y(y)$.

Example

Suppose $X \sim N(\mu, \sigma^2)$ and $Y = aX + b$ where $a > 0$. Find $f_Y(y)$.

Solution:

$$
\begin{aligned}
F_Y(y) &= P(Y \le y) \\
&= \\
&= \\
&= \qquad\qquad .
\end{aligned}
$$

Therefore,

$$f_Y(y) = \frac{d}{dy} \qquad\qquad =$$

Thus

$$f_Y(y) = \frac{1}{a\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{y-b-a\mu}{a\sigma}\right)^2\right].$$

Proposition

In the example on the previous slide, we see that Y has a normal distribution with parameters $a\mu + b$ and $a\sigma$.

Proposition A

If $X \sim N(\mu, \sigma^2)$ and $Y = aX + b$, then $Y \sim N(a\mu + b, a^2\sigma^2)$.

This proposition allows us to compute the probability $P(x_0 < X < x_1)$ where $X \sim N(\mu, \sigma^2)$ by expressing it in terms of the probability of a standard normal distribution, which we can read of the table.

Example

Let $Z = \frac{X-\mu}{\sigma}$. Then by previous proposition, $Z \sim N(0,1)$ and

$$F_X(x) = P(X \leq x) = P\left(\qquad\qquad \right)$$

$$=$$

$$= \qquad\qquad .$$

Thus

$$P(x_0 < X < x_1) = F_X(x_1) - F_X(x_0) = \Phi\left(\frac{x_1 - \mu}{\sigma}\right) - \Phi\left(\frac{x_0 - \mu}{\sigma}\right).$$

Example B

Let $X \sim N(\mu, \sigma^2)$. Find the probability that X is less than σ away from μ. (I.e., find $P(|X - \mu| < \sigma)$.)

Solution:

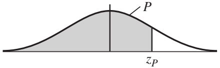
$$\begin{aligned} P(-\sigma < X - \mu < \sigma) &= P\left(-1 < \frac{X - \mu}{\sigma} < 1\right) \\ &= P(-1 < Z < 1), \end{aligned}$$

where Z follows the standard normal distribution, i.e., $Z \sim N(0, 1)$. From the table, we see that

$$P(-1 < Z < 1) = \Phi(1) - \Phi(-1) =$$

Thus $P(-\sigma < X < \sigma) = 0.6826$.
This shows that a normal random variable is within 1 standard deviation of its mean with probability .68.

TABLE 2 Cumulative Normal Distribution—Values of P Corresponding to z_p for the Normal Curve



z is the standard normal variable. The value of P for $-z_p$ equals 1 minus the value of P for $+z_p$; for example, the P for –1.62 equals 1 – .9474 = .0526.

z_p	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817

Example

Suppose X is a random variable with $X \sim N(20, 25)$. Find the probability that X satisfies $14 < X < 28$.

Solution: First, we calculate this probability by using the standard normal distribution.

$$
\begin{aligned}
P(14 < X < 28) &= \\
&= \qquad\qquad\qquad .
\end{aligned}
$$

Example C

Find the density of $X = Z^2$ where $Z \sim N(0,1)$.

Solution:

$$F_X(x) = P(X \leq x) = P(Z^2 \leq x) =$$
$$= \Phi(\sqrt{x}) - \Phi(-\sqrt{x}).$$

We find the density of X by differentiating the cdf. Note that $\Phi'(x) = \phi(x)$. By chain rule,

$$f_X(x) = \frac{d}{dx}(\Phi(\sqrt{x}) - \Phi(-\sqrt{x}))$$

$$=$$

$$=$$

$$= \qquad \text{(by symmetry of } \phi.\text{)}$$

Example C con't

From the density ϕ, we have

$$f_X(x) = \frac{x^{-1/2}}{\sqrt{2\pi}} e^{-x/2}, \qquad (x \geq 0).$$

This resembles gamma density with $\alpha = \lambda = 1/2$,

$$g(t) = \frac{t^{-1/2}}{\sqrt{2}\Gamma(\frac{1}{2})} e^{-t/2}, \qquad t \geq 0.$$

Since $f_X(x)$ and $g(t)$ only differ by at most a constant multiple, and they both integrate to 1, they must be equal, and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

This density is also called the **chi-square density** with 1 degree of freedom.

Example D

Let U be a uniform random variable on $[0, 1]$ and let $V = 1/U$. Find the density of V.

Solution:

$$F_V(v) = P(V \leq v) = P\left(\frac{1}{U} \leq v\right) = P\left(U \geq \frac{1}{v}\right)$$

$$= \qquad \qquad \text{(since U is a uniform random variable on [0,1].)}$$

Differentiating, we obtain the density

$$f_V(v) = \frac{1}{v^2}, \qquad 1 \leq v \leq \infty.$$

Proposition

Proposition B

Let X be a continuous random variable with density $f(x)$ and let $Y = g(x)$ where g is a differentiable, strictly monotonic function on some interval I. Suppose that $f(x) = 0$ if x is not in I. Then Y has the density function

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

for y such that $y = g(x)$ for some x, and $f_Y(y) = 0$ if $y \neq g(x)$ for any x in I.

Here g^{-1} is the inverse function of g; that is, $g^{-1}(y) = x$ if $y = g(x)$.

Proposition

Proposition C

Let $Z = F(X)$, that is, the cdf of X. Then Z has a uniform distribution on $[0, 1]$.

Proof:

$$
\begin{aligned}
P(Z \leq z) &= P(F(X) \leq z) \\
&= P(X \leq F^{-1}(z)) \\
&= \\
&=
\end{aligned}
$$

Proposition

Proposition D

Let U be uniform on $[0,1]$, and let $X = F^{-1}(U)$. Then the cdf of X is F.

Proof:

$$
\begin{aligned}
P(X \le x) &= P(F^{-1}(U) \le x) \\
&= \\
&= \qquad\qquad \textbf{(since } U \textbf{ is uniform on } [0,1].\textbf{)}
\end{aligned}
$$

Example E

In order to run simulations, we often need to generate random numbers. Suppose we have a uniform random number generator, apply Proposition D to generate exponential random numbers.

Solution: The cdf of $F(t) = 1 - e^{-\lambda t}$ and F^{-1} can be found by solving $x = 1 - e^{-\lambda t}$:

$$e^{-\lambda t} = 1 - x,$$

which upon taking logarithm on both sides and rearranging, we arrive at

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

Example E

Thus, if U is uniform on $[0,1]$, then $T = -\frac{1}{\lambda}\log(1-U)$ is an exponential random variable with parameter λ.

This can be simplified slightly by setting $V = 1-U$ since

$$P(V \le v) = P(1-U \le v) = P(U \ge 1-v) = (1-(1-v)) = v,$$

and so V is also uniform on [0,1].

Thus we may take $T = -\log(V)/\lambda$ where V is uniform on [0,1].