Continuous random variables

Continuous r.v.'s take an uncountably infinite number of possible values.
Examples:

- Heights of people
- Weights of apples
- Diameters of bolts
- Life lengths of light-bulbs

We cannot assign positive probabilities to every particular point in an interval. (Why?)
We can only measure intervals and their unions.

For this reason, pmf (probability mass function) is not useful for continuous r.v.'s. For them, probability distributions are described differently.

Cumulative distribution function

Defn: The (Cumulative) distribution function (or cdf) of a r.v. Y is denoted $F(y)$ and is defined as

$$F(y) = \mathbb{P}(Y \leq y)$$

for $-\infty < y < \infty$.

Example (cdf of a discrete r.v.):

Let Y be a discrete r.v. with probability function

y	1	2	3
$p(y)$	0.3	0.5	0.2

Find the cdf of Y and sketch the graph of the cdf.

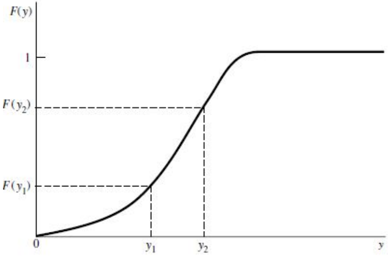
$$F(y) = \mathbb{P}(Y \leq y) = \begin{cases} 0, & \text{for } y < 1, \\ 0.3 & \text{for } 1 \leq y < 2, \\ 0.8 & \text{for } 2 \leq y < 3, \\ 1 & \text{for } y \geq 3. \end{cases}$$

Note: The cdf for any discrete r.v. is a step function.
There us a finite or countable number of points where cdf increases.

Properties of a cdf function

1. $F(-\infty) \equiv \lim_{y \to -\infty} F(y) = 0.$
2. $F(\infty) \equiv \lim_{y \to \infty} F(y) = 1.$
3. $F(y)$ is a nondecreasing function of y.
4. $F(y)$ is a right-continuous function of y.

A r.v. is said to be continuous, if its distribution function is continuous everywhere.



Probability density function

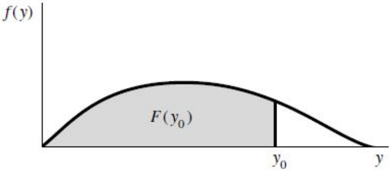
Suppose Y is a continuous r.v. with cdf $F(y)$. Assume that $F(y)$ is differentiable everywhere,
except possibly a finite number of points in every finite interval.

Then the density or probability density function (pdf) of Y is defined as

$$f(y) = \frac{dF(y)}{dy} \equiv F'(y).$$

It follows from the definitions that

$$F(y) = \int_{-\infty}^{y} f(t)dt.$$



Properties of a pdf

1. $f(y) \geq 0$ for all y
   (because $F(y)$ is non-decreasing.)
2. $\int_{-\infty}^{\infty} f(t)dt = 1$
   (because $F(\infty) = 1$.)

Example 1: Let Y be a continuous r.v. with cdf:

$$F(y) = \begin{cases} 0, & \text{for } y < 0, \\ y^3, & \text{for } 0 \leq y \leq 1, \\ 1, & \text{for } y > 1. \end{cases}$$

Find the pdf $f(y)$ and graph both $F(y)$ and $f(y)$. Find $\mathbb{P}(Y \leq 0.5)$.

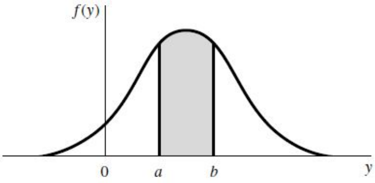
Interval probabilities

With continuous r.v.'s we usually want the probability that Y falls in a specific interval $[a, b]$; that is, we want $\mathbb{P}(a \leq Y \leq b)$ for numbers $a < b$.

Theorem If Y is a continuous r.v. with pdf $f(y)$ and $a < b$, then

$$\mathbb{P}(a \leq Y \leq b) = \int_a^b f(y)dy.$$

This represents the "area under the density curve" between $y = a$ and $y = b$:



Example

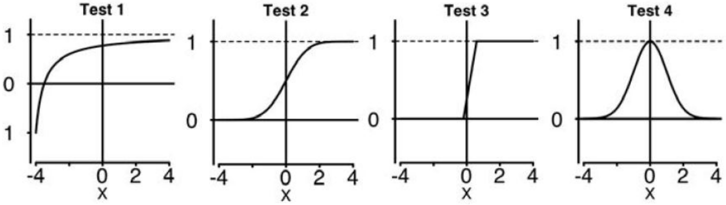
Example Let Y be a continuous r.v. with pdf:

$$f(y) = \begin{cases} 0, & \text{for } y < 0, \\ 3y^2, & \text{for } 0 \leq y \leq 1, \\ 0, & \text{for } y > 1. \end{cases}$$

Find $\mathbb{P}(0.4 \leq y \leq 0.8)$.

Quiz

Which of the following are graphs of valid cumulative distribution functions?



(A) 1, 2, and 3
(B) 1 and 2
(C) 2 and 3
(D) 3 only
(E) 4 only

Quiz

Suppose X has range $[0,2]$ and pdf $f(x) = cx$.

What is the value of c?

(A) 1/4
(B) 1/2
(C) 1
(D) 2
(E) Other

Quiz

Suppose X has range $[0, 2]$ and pdf $f(x) = cx$.

Compute $\mathbb{P}(1 \leq X \leq 2)$.

(A) 1/4
(B) 1/2
(C) 3/4
(D) Other

Quiz

Suppose X has range $[0, b]$ and cdf $f(x) = x^2/9$.

What is b?

Quiz

Suppose X is a continuous random variable.

What is $\mathbb{P}(a \leq X \leq a)$?

Quiz

Suppose X is a continuous random variable.

Does $\mathbb{P}(X = a) = 0$ mean X never equals a?

(A) Yes
(B) No

What does it mean never happens?

The number of possible deals in bridge card game is

$$\binom{52}{13, 13, 13, 13} = \frac{52!}{13!13!13!13!} \approx 5 \times 10^{28}$$

. If we try and deal cards every second until we find a specific combination, then the time until success is a geometric random variable with the expected time equal to 5×10^{28} seconds.

The age of the Universe is currently estimated as 13.7×10^9 years. Every year has $60 \times 60 \times 24 \times 365 \approx 31 \times 10^6$ seconds. So the age of the Universe is $\approx 4 \times 10^17$ seconds

So the time to wait for a specific combination of cards in the bridge game is significantly longer than the age of the Universe.

Why continuous r.v. take values?

Why can't we argue as below:

$$\mathbb{P}(X \in [0,1] = \mathbb{P}(\bigcup_{0 \le a \le 1} \{X = a\}) = \sum_{0 \le a \le 1} \mathbb{P}(X = a) = \sum_{0 \le a \le 1} 0 = 0 \ ?$$

Axioms of probability theory do not allow us to speculate
about the probability of a union of the uncountable number of events.

Example

Let Y be a continuous r.v. with pdf

$$f(y) = \begin{cases} cy^{-1/2}, & \text{for } 0 < y \leq 4, \\ 0, & \text{elsewhere,} \end{cases}$$

for some constant c.

Find c, find $F(y)$, graph $f(y)$ and $F(y)$, and find $\mathbb{P}(Y < 1)$.

Expected value of a continuous random variable

Defn: The expected value of a continuous r.v. Y is

$$\mathbb{E}(Y) = \int_{-\infty}^{\infty} yf(y)dy,$$

provided that the integral exists.

Theorem If $g(Y)$ is a function of a continuous r.v. Y, then

$$\mathbb{E}[g(Y)] = \int_{-\infty}^{\infty} g(y)f(y)dy.$$

This is similar to the case of a discrete r.v., with an integral replacing a sum.

Properties of the expected value

- $\mathbb{E}[aY + b] = a\mathbb{E}Y + b$.
- $\mathbb{E}[Y_1 + Y_2 + \ldots + Y_n] = \mathbb{E}Y_1 + \mathbb{E}Y_2 + \ldots + \mathbb{E}Y_n$.

The variance and the standard deviation are defined in the same way as for the discrete r.v.:

$$\mathrm{Var}(Y) = \mathbb{E}[(Y - \mu)^2] = \mathbb{E}[Y^2] - \mu^2,$$

where $\mu = \mathbb{E}Y$.

And

$$\sigma = \sqrt{\mathrm{Var}(Y)}.$$

Example

Find the expectation and variance for the r.v. in the previous example.

Pdf is

$$f(y) = \begin{cases} cy^{-1/2}, & \text{for } 0 < y \leq 4, \\ 0, & \text{elsewhere,} \end{cases}$$

for some constant c.

Quiz

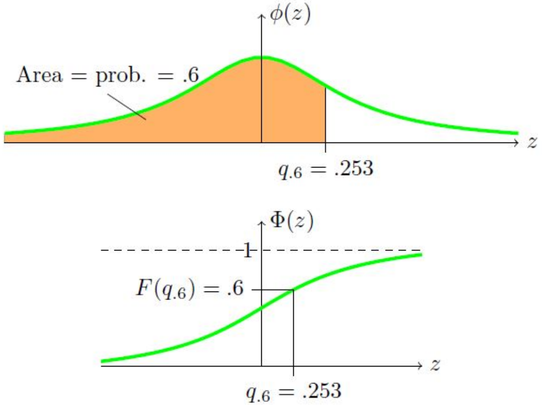
True or False. The mean divides the probability mass in half.

(A) True
(B) False

Quantiles

Defn: Let $0 < x < 1$. The x-th quantile of a r.v. Y (denoted q_x) is the smallest value such that $\mathbb{P}(Y \leq q_x) \geq x$.

If Y is continuous, q_x is the smallest value such that $\mathbb{P}(Y \leq q_x) = x$.



The quantile function is the inverse of the cumulative distribution function.

Median

The 50% quantile of a random variable Y is called the median of Y.

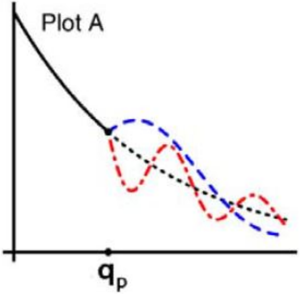
Example: What is the median in the previous example?

In R, quantiles can be computed by functions like *qbinom* or *qpois*.

Example: what is the median of the Poisson distribution with parameter $\lambda = 7$.

Quiz

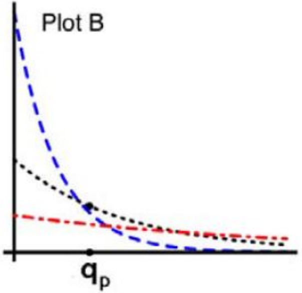
In the plot below some densities are shown. The median of the black plot is always at q_p. Which density has the greatest median?



(A) Black
(B) Red
(C) Blue
(D) All the same
(E) Impossible to tell

Quiz

In the plot below some densities are shown. The median of the black plot is always at q_p. Which density has the greatest median?

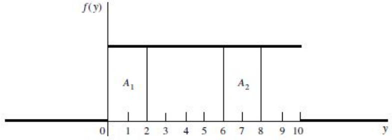


(A) Black
(B) Red
(C) Blue
(D) All the same
(E) Impossible to tell

The Uniform Distribution

Example Suppose a bus is supposed to arrive at a bus stop at 8 : 00 AM. In reality, it is equally likely to arrive at any time between 8 : 00 AM and 8 : 10 AM.

Let Y be the amount of time between 8 : 00 AM and the moment when the bus arrives. This is a simple example of a continuous r.v.



What is the probability the bus will arrive between 8 : 02 and 8 : 06 AM?

The uniform distribution

Defn: A r.v. Y has a uniform probability distribution on the interval $[\theta_1, \theta_2]$ if its pdf is

$$f(y) = \begin{cases} \frac{1}{\theta_2 - \theta_1}, & \text{if } \theta_1 \leq y \leq \theta_2 \\ 0, & \text{otherwise.} \end{cases}$$

(If parameters $\theta_1 = 0$ and $\theta_2 = 1$, the distribution is called the standard uniform distribution.)

Theorem If $Y \sim \mathrm{Unif}(\theta_1, \theta_2)$, then

$$\begin{aligned} \mathbb{E}(Y) &= \frac{\theta_1 + \theta_2}{2}, \\ \mathrm{Var}(Y) &= \frac{(\theta_2 - \theta_1)^2}{12}. \end{aligned}$$

Proof:

Quiz

Find the variance of $X \sim \mathrm{Unif}(0, 4)$.

(A) 1/12
(B) 1/3
(C) 4/3
(D) 2
(E) other

Quiz

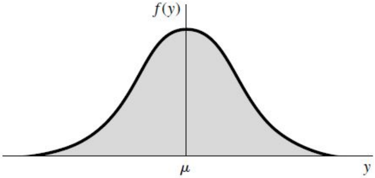
Find the variance of the sum of four independent $\mathrm{Unif}(0, 4)$ random variables.

Quiz

Find the 0.7 quantile of $X \sim \mathrm{Unif}(2, 32)$.

(A) 21
(B) 23
(C) 24
(D) 25
(E) other

The normal (Gaussian) distribution



The normal distribution is called normal because it describes well many real-world random variables. It is also often called Gaussian.

Defn: A r.v. Y has a normal or Gaussian distribution, if its pdf is

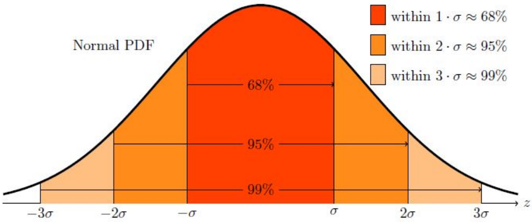
$$f(y) = \frac{1}{\sqrt{2\pi\sigma^2}} exp^{-(y-\mu)^2/(2\sigma^2)}.$$

Notation: $Y \sim \mathcal{N}(\mu, \sigma^2)$.

The normal pdf is defined over entire real line.

Standard normal distribution

If $\mu = 0$ and $\sigma = 1$, then the distribution is called the standard normal distribution.



Rules of thumb:
$\mathbb{P}(-1 \leq Z \leq 1) \approx .68$,
$\mathbb{P}(-2 \leq Z \leq 2) \approx .95$,
$\mathbb{P}(-3 \leq Z \leq 3) \approx .997$.

Quiz

Suppose a p-quantile of the standard normal distribution equals 2. To which of the following is p closest to?

(A) 99.7%
(B) 97.5%
(C) 95%
(D) 5%
(E) 2.5 %

Normal distribution as an approximation

Let $X_1, X_2, \dots X_n$ be independent identically distributed r.v. with expectation μ and variance σ^2.

Then

$$Y = \frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sqrt{n\sigma^2}}$$

is distributed approximately as a standard normal r.v.

R-simulation

Generating data:
repet <- 10000
size <- 100
p <- .5
data <- (rbinom(repet, size, p) - size * p) / sqrt(size * p * (1-p))

Histograms:

- By default you get a histogram with the counts of items in each bin
  To get a density you set freq = FALSE
- controlling the number of bins:
  'breaks = a number' sets the number of bins to use in the histogram

hist(data, breaks = size/2, col = 'red', freq = FALSE)

Plotting the pdf of the normal distribution:

x = seq(min(data) - 1, max(data) + 1, .01)
lines(x, dnorm(x), col='green', lwd = 4)

Properties of normal distribution

The function

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

is a valid density function:

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

Proof: First let us change variable, $t = (x-\mu)/\sigma$.

The Gaussian integral

We need to compute

$$I = \int_{-\infty}^{\infty} e^{-t^2/2} dt$$

Let us try a change of variable: $x = \frac{t^2}{2}$, $t = \sqrt{2x}$, $dt = \frac{1}{\sqrt{2}} x^{-1/2} dx$.
Then we get

$$I = \sqrt{2} \int_0^{\infty} x^{-1/2} e^{-x} dx.$$

The integral $\int_{-\infty}^{\infty} x^{a-1} e^{-x} dx$ is called the Gamma function $\Gamma(a)$ and therefore

$$I = \sqrt{2}\Gamma\left(\frac{1}{2}\right).$$

It is easy to calculate that $\Gamma(n) = (n-1)!$ for positive integers n.
But what is $\Gamma\left(\frac{1}{2}\right)$?

A trick

$$I^2 = \int_{-\infty}^{\infty} e^{-x^2/2} \int_{-\infty}^{\infty} e^{-y^2/2} dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)/2} dxdy.$$

Let us use the polar coordinates:

$$I^2 = \int_0^{2\pi} \int_0^{\infty} re^{-r^2/2} drd\theta$$

$$I^2 = -2\pi \ e^{-r^2/2}\Big|_0^{\infty} = 2\pi.$$

So $I = \sqrt{2\pi}$ and $\frac{1}{\sqrt{2\pi}} e^{-t^2/2}$ is a density.

And $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, which is surprising.

Standartization

Theorem. If $Y \sim N(\mu, \sigma^2)$, then the standardized version of Y,

$$Z = \frac{Y - \mu}{\sigma},$$

has the standard normal distribution $N(0, 1)$.

Proof: Let us calculate the cdf of Z:

$$F_Z(t) = \mathbb{P}(Z \leq t) = \mathbb{P}(\frac{Y - \mu}{\sigma} \leq t) = \ldots$$

.

Mean and variance of normal distribution

Theorem If $Y \sim N(\mu, \sigma)$, then

$$\mathbb{E}(Y) = \mu,$$

$$\mathrm{Var}(Y) = \sigma^2.$$

Proof: Since $Y = \mu + \sigma Z$, where Z is the standard normal r.v., let us first compute $\mathbb{E}Z$ and $\mathrm{Var}Z$.

Mgf of normal distribution

Theorem If $Z \sim N(0,1)$, then its moment-generating function is

$$m_Z(t) = e^{t^2/2}.$$

Proof:

Similarly, the moment generating function of $Y \sim N(\mu, \sigma^2)$ is $m_Y(t) = e^{\mu t + \frac{\sigma^2}{2} t^2}$.

Consequence: The odd moments of the standard normal variable are zero and the even moments are:
$m'_2 = 1$, $m'_3 = 0$,
$m'_4 = 3$, $m'_5 = 0$,
$m'_6 = 3 \cdot 5$, $m'_7 = 0$,
$m'_8 = 3 \cdot 5 \cdot 7$,
...
$m'_{2n} = (2n-1)!! = \frac{1}{\sqrt{\pi}} \Gamma\left(n + \frac{1}{2}\right)$

Calculating normal probabilities

If $Y \sim N(\mu, \sigma^2)$, the cdf for Y is

$$\mathbb{P}(Y \leq y) \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi\sigma^2}} exp^{-(x-\mu)^2/(2\sigma^2)} dx = 1.$$

However, this integral cannot be expressed in elementary functions.

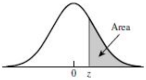
It can be calculated via numerical integration.

We use software or tables to calculate such normal probabilities.

Example

Example. If $Z \sim N(0,1)$, find
$\mathbb{P}(Z > 1.83)$, $\mathbb{P}(Z > -0.42)$, $\mathbb{P}(Z \leq 1.19)$, $\mathbb{P}(-1.26 < Z \leq 0.35)$

Table 4 Normal Curve Areas
Standard normal probability in right-hand tail
(for negative values of z, areas are found by symmetry)



	Second decimal place of z									
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064

Example

Using standartization, we can use Table 4 to find probabilities for any normal r.v.

Example A graduating class has GPAs that follow a normal distribution with mean 2.70 and variance 0.16.
What is the probability a randomly chosen student has a GPA greater than 2.50?

What is the probability that a random GPA is between 3.00 and 3.50?

Exactly 5% of students have GPA above what number?

Example

Example: Let Y be a normal r.v. with $\sigma = 10$. Find μ such that $P(Y < 10) = 0.75$.

Quiz

The length of human gestation (pregnancy) is well-approximated by a normal distribution with mean $\mu = 280$ days and standard deviation $\sigma = 8.5$ days. Suppose your final exam is scheduled for May 18 and your pregnant professor has a due date of May 25.

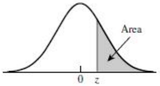
Find the probability she will give birth on or before the day of the final.

Quiz

Find the probability she will give birth on or before the day of the final.

(A) ≈ 2.5% (B) ≈ 5% (C) ≈ 10% (D) ≈ 20% (E) ≈ 30%

Table 4 Normal Curve Areas
Standard normal probability in right-hand tail
(for negative values of z, areas are found by symmetry)



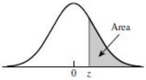
	Second decimal place of z									
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143

Quiz

The professor decides to move up the exam date so there will be a 95% probability that she will give birth afterward. What date should she pick?

(A) May 20 (B) May 15 (C) May 11 (D) May 8 (E) May 6

Table 4 Normal Curve Areas
Standard normal probability in right-hand tail
(for negative values of z, areas are found by symmetry)



	Second decimal place of z									
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110

The Gamma distribution and related distributions

Many variables have distributions that are non-negative and skewed to the right (long right tail).

Examples:

- Lifelengths of manufactured parts
- Lengths of time between arrivals at a restorant
- Survival times for severely ill patients

Such r.v.'s may often be modeled with a Gamma distribution.

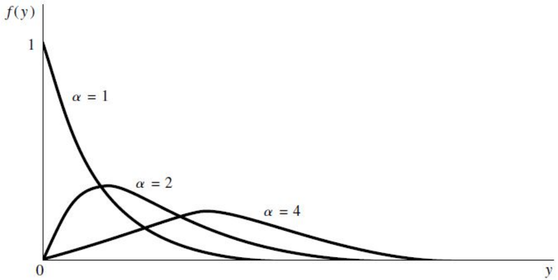
The Gamma distribution

Defn: A continuous r.v. has a Gamma distribution if its pdf is

$$f(y) = \frac{1}{\beta^\alpha \Gamma(\alpha)} y^{\alpha-1} e^{-y/\beta},$$

where $\alpha > 0$, $\beta > 0$, and $y \in [0, \infty)$.

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



α and β are called the shape and scale parameters, respectively.

The mean and variance of gamma distribution

Theorem: If Y has a gamma distribution with parameters α and β, then

$$\mathbb{E}(Y) = \alpha\beta,$$

$$\mathrm{Var}(Y) = \alpha\beta^2.$$

Proof:

The mgf of gamma distribution

Theorem: If Y has a gamma distribution with parameters α and β, then its moment generating function is

$$m_Y(t) = \frac{1}{(1-\beta t)^\alpha}.$$

Proof:

Example

Let $f(y) = cy^4 e^{-y/3}$ if $y \geq 0$ and zero elsewhere.
What value of c makes $f(y)$ a valid density?

What is $E(Y)$? What is $V(Y)$

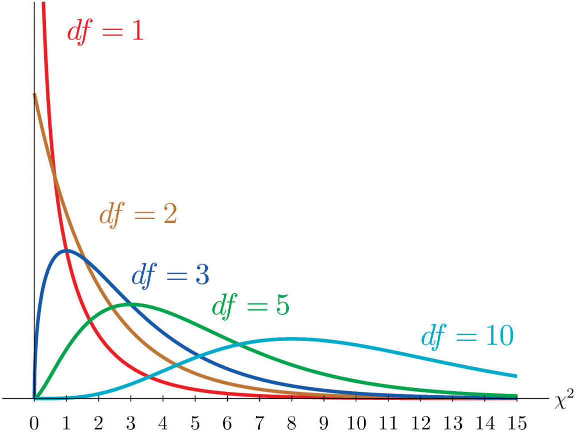
Example

Suppose $Y \sim \Gamma(5, 3)$. Find the probability that Y falls within 2 standard deviations of its mean.

First special case: χ^2 distribution.

Defn: For any integer $\nu \geq 1$, a r.v. Y has a chi-square (χ^2) distribution with ν degrees of freedom if Y is a Gamma r.v. with $\alpha = \nu/2$ and $\beta = 2$.

[Shorthand: $Y \sim \chi^2(\nu)$]



In other words, χ^2 distribution is the Gamma distribution for half-integer α and a specific choice of β.

Properties of χ^2 distribution

If $X_1, X_2, \ldots X_\nu$ are independent standard normal r.v.'s, then $Y = (X_1)^2 + (X_2)^2 + \ldots (X_\nu)^2$ has the χ^2 distribution with ν degrees of freedom.

Proof:

χ^2 distribution often occurs in statistics.

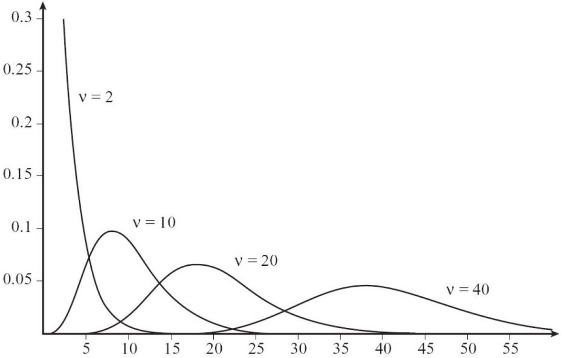
Mean, variance and mgf of χ^2 distribution

Theorem: If $Y \sim \chi^2(\nu)$, then

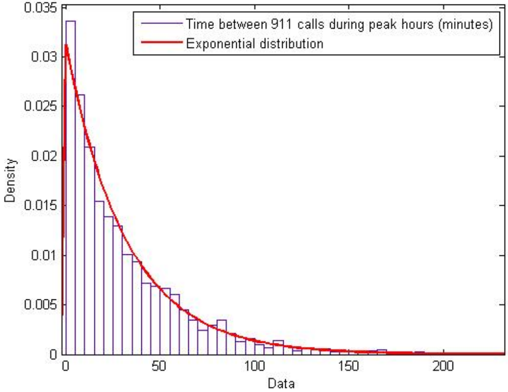
$$E(Y) = \nu,$$

$$V(Y) = 2\nu,$$

$$m_Y(t) = \frac{1}{(1-2t)^{\nu/2}}.$$



Second special case: Exponential distribution.



The exponential distribution is commonly used as a model for lifetimes, survival times, or waiting times.

Mean, variance, and mgf of exponential distribution.

Defn: A r.v. Y has an exponential distribution with parameter β if Y is a Gamma r.v. with $\alpha = 1$. That is, the pdf of Y is

$$f(y) = \frac{1}{\beta} e^{-y/\beta}$$

for $y \geq 0$.

Theorem If Y is an exponential r.v. with parameter β, then

$$\mathbb{E}(Y) = \beta,$$

$$\mathrm{Var}(Y) = \beta^2.$$

$$m_Y(t) = \frac{1}{1 - \beta t}$$

Example

Let the lifetime of a part (in thousands of hours) follow an exponential distribution with mean lifetime 2000 hours. ($\beta = \ldots$)

Find the probability the part lasts less than 1500 hours.

Find the probability the part lasts more 2200 hours.

Find the probability the part lasts between 2000 and 2500 hours.

Memoryless property of exponential distribution

Let a lifetime Y follow an exponential distribution. Suppose the part has lasted a units of time already. Then the conditional probability of it lasting at least b additional units of time is

$$\mathbb{P}(Y > a + b | Y > a)$$

.

It is the same as the probability of lasting at least b units of time in the first place:

$$\mathbb{P}(Y > b)$$

Proof:

The exponential distribution is the only continuous distribution with this "memoryless"property.
It is the continuous analog of the geometric distribution.

Relationship with Poisson process

Suppose we have a Poisson process X with a mean of λ events per time unit. Consider the waiting time W until the first occurrence. Then, $W \sim \text{Exp}(1/\lambda)$.

Proof:

$$\mathbb{P}(W \leq t) = \mathbb{P}(X_{[0,t]} \geq 1) = 1 - \mathbb{P}(X_{[0,t]} = 0) = \dots$$

.

This result can be generalized: For any integer $\alpha \geq 1$, the waiting time W until the α-th occurence has a gamma distribution with parameters α and $\beta = 1/\gamma$.

Example

Suppose customers arrive in a queue according to a Poisson process with mean $\lambda = 12$ per hour. What is the probability we must wait more than 10 minutes until the first customer?

What is the probability we must wait more than 10 minutes to see the fourth customer?

Beta distribution

Suppose we want to model a random quantity that takes values between 0 and 1.

Examples:

- The proportion of a chemical product that is pure.
- The proportion of a hospital's patients infected with a certain virus.
- The parameter p of a binomial distribution.

We can model this random quantity with a beta-distributed random variable.

More generally, if a random quantity Y has its support on interval $[c, d]$ then we can model the transformed quantity $Y^* = \frac{Y-c}{d-c}$ with a beta r.v.

Beta distribution

Defn: A r.v. Y has a beta distribution with parameters $\alpha > 0$ and $\beta > 0$ [Shorthand: $Y \sim Beta(\alpha, \beta)$] if its pdf is:

$$f(y) = \frac{1}{B(\alpha, \beta)} y^{\alpha-1}(1-y)^{\beta-1},$$

where

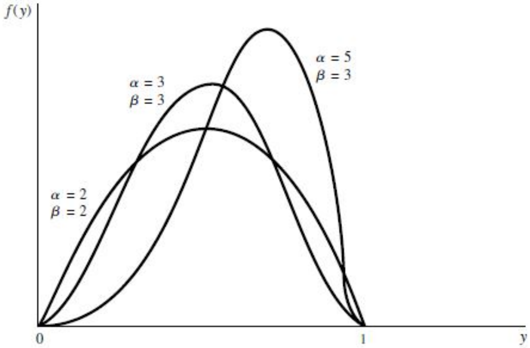
$$B(\alpha, \beta) = \int_0^1 y^{\alpha-1}(1-y)^{\beta-1} dy = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Note 1: The beta pdf has support on $[0, 1]$.

Note 2: For integer α and β,

$$\frac{1}{B(\alpha, \beta)} = \frac{(\alpha+\beta-1)!}{(\alpha-1)!(\beta-1)!}.$$

Graphs of beta distribution



Beta distribution is quite flexible.

Quiz

What is another name for Beta$(1, 1)$?

(A) Normal
(B) Exponential
(C) Poisson
(D) Chi-square
(E) Uniform

Mean and variance of beta distribution

Theorem: If $Y \sim \textit{Beta}(\alpha, \beta)$, then

$$\mathbb{E}(Y) = \frac{\alpha}{\alpha + \beta},$$

$$\mathrm{Var}(Y) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

What happens if α and β grow ?

Relation to Bayesian Statistics

Suppose we perform an experiment with outcomes F and S. The probability of outcome S is p.

We model p as a r.v. and our initial belief about the distribution of p is that it is distributed according to Beta distribution with parameters α and β.

We observed outcome S. How we should update our beliefs?

The posterior distribution of p is Beta$(\alpha + 1, \beta)$.

What happens with the distribution if we repeat the experiment 100 times with 90 successes and 10 failures?

Finding beta probabilities

If α and β are integers, then one can use the formula

$$F(y) = \frac{1}{B(\alpha, \beta)} \int_0^y t^{\alpha-1}(1-t)^{\beta-1}dt = \sum_{i=\alpha}^{n} \binom{n}{i} y^i (1-y)^{n-i},$$

where $n = \alpha + \beta - 1$. This formula can be established by partial integration.

That is, we can write

$$F(y) = \mathbb{P}(X \geq \alpha) = 1 - \mathbb{P}(X \leq \alpha - 1),$$

where $X \sim \text{Binom}(n, y)$.

In general, one can use R functions like $pbeta(y_0, \alpha, 1/\beta)$ for cdf and $qbeta(p, \alpha, 1/\beta)$ for quantiles.

Example

Define the downtime rate as the proportion of time a machine is under repair. Suppose a factory produces machines whose downtime rate follows a $beta(3, 18)$ distribution.

What is the pdf of the downtime rate Y?

For a randomly selected machine, what is the expected downtime rate?

What is the probability that a randomly selected machine is under repair less than 5% of the time?

If machines act independently, in a shipment of 25 machines, what is the probability that at least 3 have downtime rates greater than 0.20?

Mixed distributions

A mixed distribution has probability mass placed at one or more discrete points, and has its remaining probability spread over intervals.

Example Let Y be the annual amount paid out to a random customer by an auto insurance company.

Suppose $3/4$ of customers each year do not receive any payout.

Of those who do receive benefits, the payout amount Y_2 follows a gamma(10,160) distribution.

For a random customer, what is the expected payout amount?

CDF of a mixed distribution

If Y has a mixed distribution, then the cdf of Y may be written as

$$F(y) = c_1 F_1(y) + c_2 F_2(y),$$

where $F_1(y)$ is a step distribution function, $F_2(y)$ is a continuous distribution function, c_1 is the total mass of all discrete points, and c_2 is the total probability of all continuous portions.
We can think about the distributions $F_1(y)$ and $F_2(y)$ as cdf's of two r.v.'s, X_1 (discrete) and X_2 (continuous).
What are c_1, c_2, mpf of X_1 and *pdf* of X_2 in the previous example?

For any function $g(Y)$ of Y,

$$\mathbb{E}g(Y) = c_1\mathbb{E}g(X_1) + c_2\mathbb{E}g(X_2).$$

For a random customer, what is the expected payout amount?

Example

In a lab experiment, a rat is given a shot of a dangerous toxin. There is a probability 0.2 that the rat will die instantly. If it survives the initial shot, suppose it would ordinarily then die at a random point in time over the next 24 hours. However, if it is still alive 6 hours after the shot, the rat is killed.

Find the expected value of the survival time for a rat entering this experiment.
What is the probability that a rat entering the experiment will survive 3 hours or less?

Chebyshev's theorem one more time

Theorem: Let Y be a r.v. with mean μ and variance σ^2. Then for every $k > 0$, we have

$$\mathbb{P}(|Y-\mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

Proof:
Let $X = (Y-\mu)^2/(k\sigma)^2$. We need to prove that $\mathbb{P}(X \geq 1) \leq 1/k^2$.
Define the indicator function of the event $A = \{X \geq 1\}$:

$$1_A(x) = \begin{cases} 1, & \text{if } x \geq 1, \\ 0, & \text{otherwise.} \end{cases}$$

Then $\mathbb{P}(X \geq 1) = \mathbb{E}[1_A(X)]$.
Note that $1_A(X) \leq X$,
hence $\mathbb{P}(X \geq 1) = \mathbb{E}[1_A(X)] \leq \mathbb{E}X = \mathbb{E}[(Y-\mu)^2/(k\sigma)^2] = 1/k^2$. $\square$