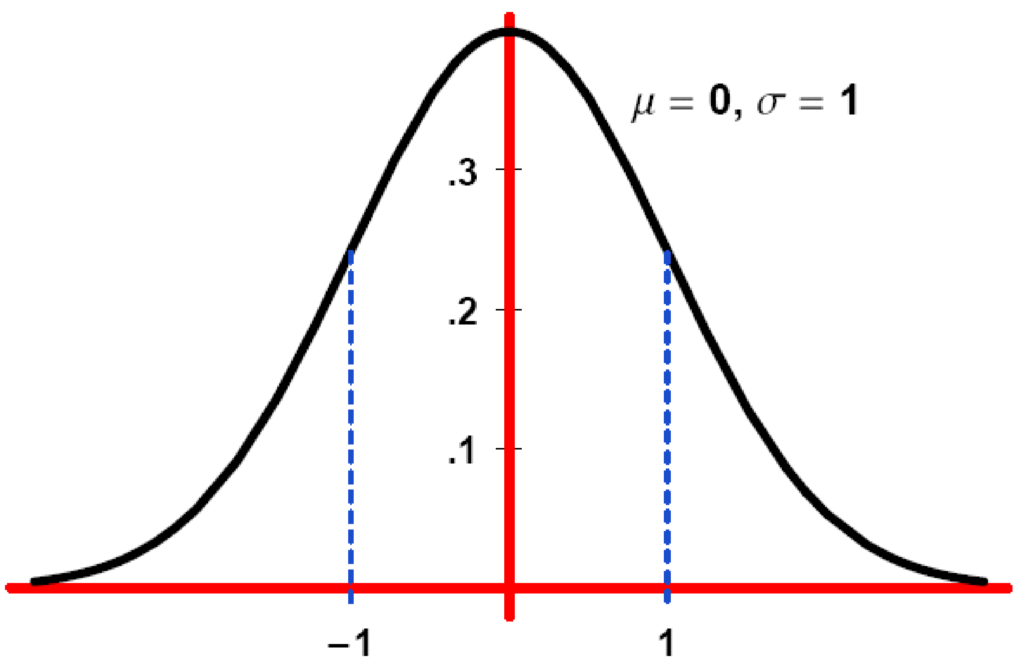
SOME SPECIFIC PROBABILITY DISTRIBUTIONS

1. Normal random variables

1.1. Probability Density Function. The random variable X is said to be normally distributed with mean μ and variance σ^2 (abbreviated by $x \sim N[\mu, \sigma^2]$ if the density function of x is given by

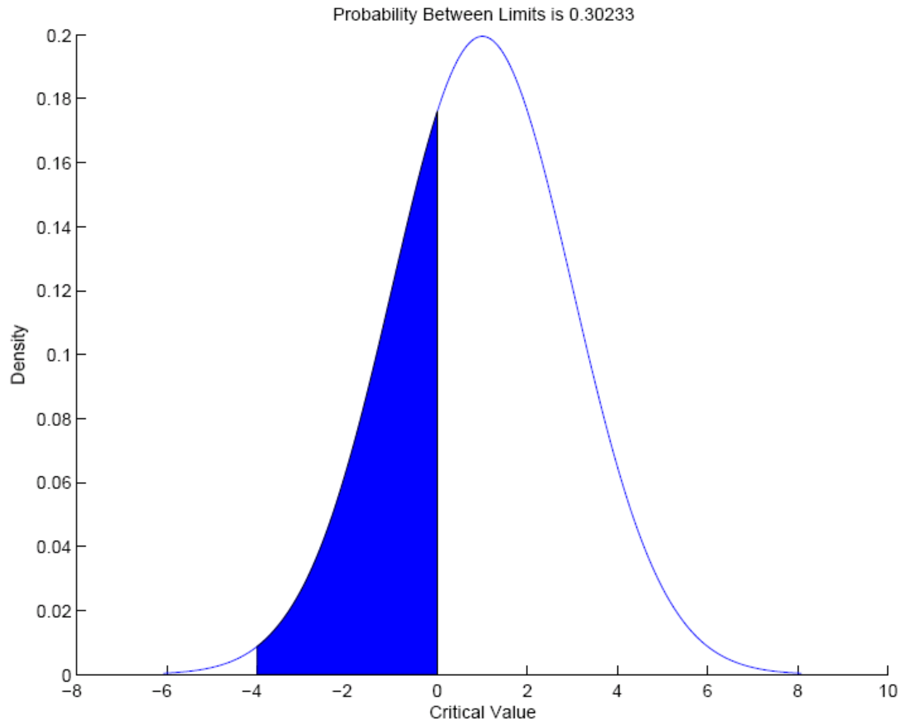
$$f\,(x\,;\;\mu, \sigma^2)\;=\;\frac{1}{\sqrt{2\pi\sigma^2}}\cdot e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^{\,2}}$$

The normal probability density function is bell-shaped and symmetric. The figure below shows the probability distribution function for the normal distribution with a $\mu = 0$ and $\sigma = 1$. The areas between the two lines is 0.68269. This represents the probability that an observation lies within one standard deviation of the mean.



Date: September 8, 2006.

The next figure shows the portion of the distribution between -4 and 0 when the mean is one and σ is equal to two.



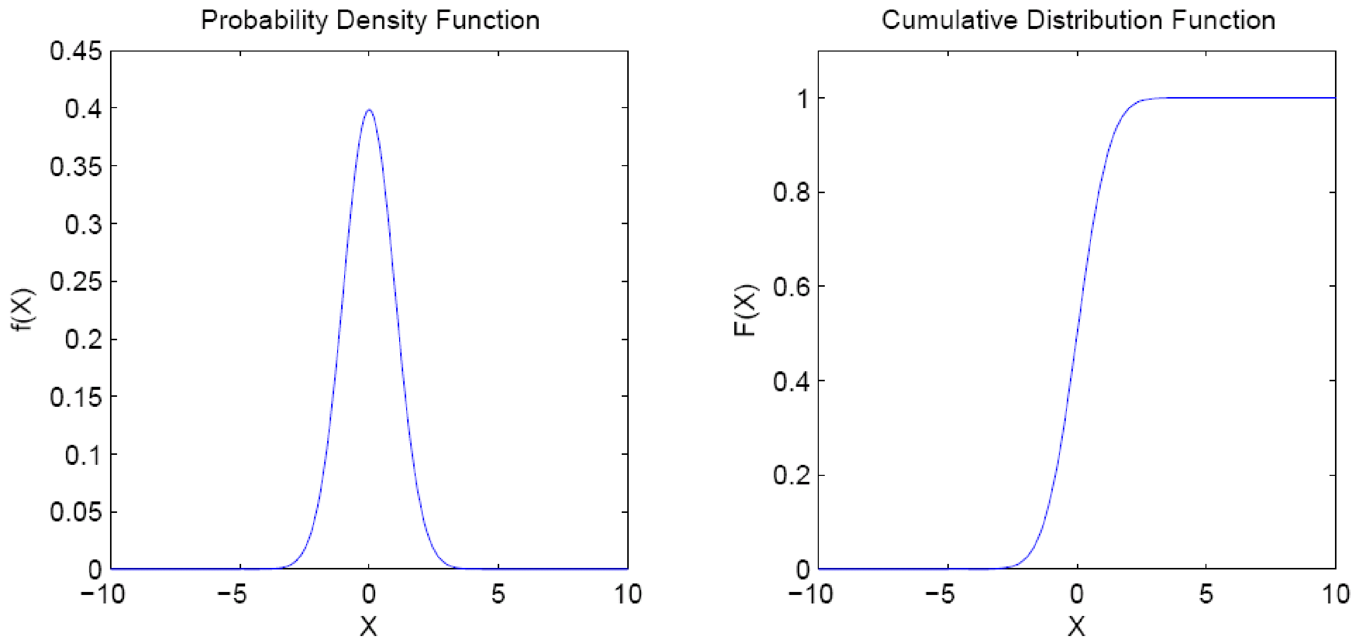
1.2. **Properties of the normal random variable.**

a: $E(x) = \mu$, $Var(x) = \sigma^2$.
b: The density is continuous and symmetric about μ.
c: The population mean, median, and mode coincide.
d: The range is unbounded.
e: There are points of inflection at $\mu \pm \sigma$.
f: It is completely specified by the two parameters μ and $\sigma^{\,2}$.
g: The sum of two independently distributed normal random variables is normally distributed.
If $Y = \alpha X_1 + \beta X_2 + \gamma$ where $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ and X_1 and X_2 are independent, then $Y \sim N(\alpha\mu_1 + \beta\mu_2 + \gamma; \alpha^2\sigma_1^2 + \beta^2\sigma_2^2)$.

1.3. **Distribution function of a normal random variable.**

$$F(x\,;\ \mu,\, \sigma^2) \;=\; Pr\,(X \;\leq\; x) \;=\; \int_{-\infty}^{x} f\,(s\,;\, \mu,\, \sigma^2\,)ds$$

Here is the probability density function and the cumulative distribution of the normal distribution with $\mu = 0$ and $\sigma = 1$.

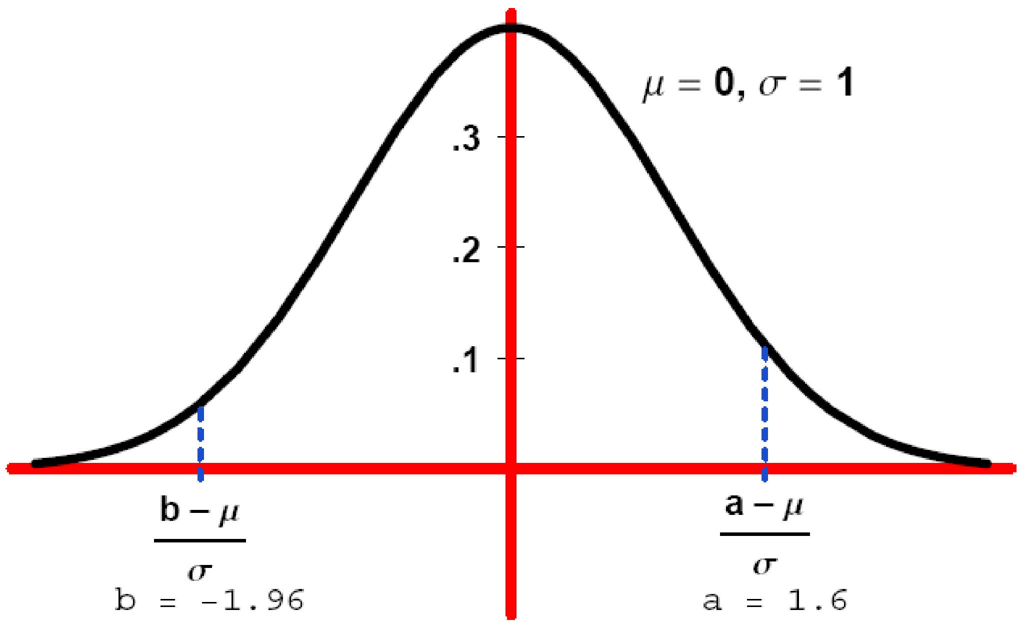


1.4. **Evaluating probability statements with a normal random variable.** If $x \sim N(\mu, \sigma^2)$ then,

$$\begin{aligned} Z &= \tfrac{X-\mu}{\sigma} \;\sim\; N(0,1) \\ E\left(Z\right) &= E\left(\tfrac{X-\mu}{\sigma}\right) \;=\; \tfrac{1}{\sigma}\cdot\left(E(X)-\mu\right) \;=\; 0 \\ Var\left(Z\right) &= Var\left(\tfrac{X-\mu}{\sigma}\right) \;=\; \tfrac{1}{\sigma^2}Var(X-\sigma) \\ &= \tfrac{\sigma^2}{\sigma^2} \;=\; 1 \end{aligned}$$

Consequently,

$$\begin{aligned} Pr(a \le x \le b) &= Pr\left(a - \mu \le x - \mu \le b - \mu\right) \\ &= Pr\left[\frac{a-\mu}{\sigma} \le \frac{x-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right] \\ &= F\left(\frac{b-\mu}{\sigma};\ 0,\ 1\right) - F\ \left(\frac{a-\mu}{\sigma};\ 0,\ 1\right) \\ &= area\ below \end{aligned}$$



We can then merely look in tables for the distribution function of a N(0,1) variable.

1.5. **Moments of a normal random variable.** The first central moment is

$$\begin{aligned} E\left(X-\mu\right) &= \int_{-\infty}^{\infty} (x-\mu)\,\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx \\ &= \int_{-\infty}^{0} (x-\mu)\,\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx + \int_{0}^{\infty} (x-\mu)\,\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx \\ &= -\int_{0}^{\infty} (x-\mu)\,\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx + \int_{0}^{\infty} (x-\mu)\,\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx \\ &= 0 \end{aligned}$$

because $(x-\mu)\,\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$ is an odd function.

The second central moment is

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

Use the following for integration by parts:

$$\begin{aligned} dv &= (x-\mu)\, e^{-\frac{1}{2\sigma^2}(x-\mu)^2},\ v = -\sigma^2 e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \\ u &= (x-\mu)\,,\ du = dx \end{aligned}$$

Then we get:

$$\begin{aligned} E(X - \mu)^2 &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (x-\mu)^2\, e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx \\ &= -\frac{1}{\sqrt{2\pi\sigma^2}} \sigma^2 (x-\mu)\, e^{-\frac{1}{2\sigma^2}(x-\mu)^2} |_{-\infty}^{\infty} - \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \left(-\sigma^2 e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \right) dx \\ &= \sigma^2 \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx = \sigma^2 \end{aligned}$$

The third central moment is 0 because $(x-\mu)^3 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$ is an odd function.

The first raw moment is μ. This follows from

$$\begin{aligned} E\,(X-\mu) &= 0 \Leftrightarrow \\ E\,(X) - \mu &= 0 \Leftrightarrow \\ E\,(X) &= \mu \end{aligned}$$

The secon raw moment can be found using the secon central moment:

$$\begin{aligned} E\,(X-\mu)^2 &= \sigma^2 \Leftrightarrow \\ E\left(X^2\right) - 2\mu E\,(X) + \mu^2 &= \sigma^2 \Leftrightarrow \\ E\left(X^2\right) - \mu^2 &= \sigma^2 \Leftrightarrow \\ E\left(X^2\right) &= \sigma^2 + \mu^2 \end{aligned}$$

2. Chi-square random variable

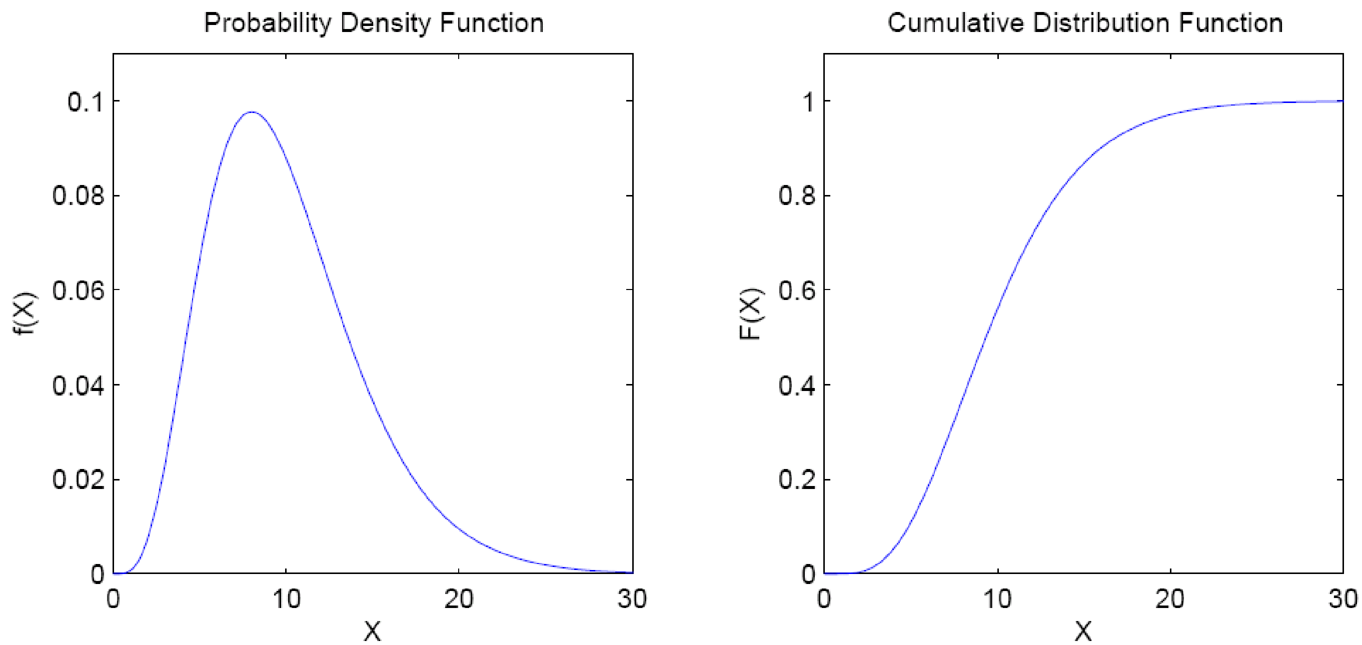
2.1. Probability Density Function. The random variable X is said to be a chi-square random variable with ν degrees of freedom [abbreviated $\chi^2(\nu)$] if the density function of X is given by

$$f(x;\nu) = \begin{cases} \frac{1}{2^{\frac{\nu}{2}}\,\Gamma\left(\frac{v}{2}\right)}\, x^{\frac{\nu}{2}-1} e^{\frac{-x}{2}} & 0 < x \\ 0 & otherwise \end{cases}$$

where $\Gamma(\cdot)$ is the gamma function defined by

$$\Gamma(r) = \int_0^{\infty} u^{r-1}\, e^{-u}\, du, \quad r > 0$$

Note that for positive integer values of r, $\Gamma(r) = (r-1)!$



The following diagram shows the pdf and cdf for the chi-square distribution with parameters $\nu = 10$.

2.2. **Properties of the chi-square random variable.**

2.2.1. χ^2 *and* $N(0,1)$. Consider n independent random variables.

$$X_i \sim N(0,1),\ i = 1, 2, ..., n$$
$$then \sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$$

It can also be shown that

$$If\ X_i \sim N(0,1),\ i = 1, 2, ..., n$$
$$and\ \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$
$$then \sum_{i=1}^{n-1} \left(X_i - \bar{X}\right)^2 \sim \chi^2(n)$$

because this is the sum of $(n-1)$ independent random variables given that $\bar{X}$ and $(n-1)$ of the x's are independent.

2.2.2. χ^{2} $andN(\mu, \sigma^{2})$.

$$If\ X_i \sim N(\mu, \sigma^2),\ i = 1, 2, ..., n$$

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

$$then\ \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^{2}(n)$$

$$and\ \sum_{i=1}^{n-1} \left(\frac{X_i - \bar{X}}{\sigma} \right)^2 \sim \chi^{2}(n)$$

2.2.3. *Sums of chi-square random variables.* If y_1 and y_2 are independently distributed as $\chi^2(\nu_1)$ and $\chi^2(\nu_2)$, respectively, then

$$y_1 + y_2 \sim \chi^2(\nu_1 + \nu_2).$$

2.2.4. *Moments of chi-square random variables.*

$$Mean\ \left(\chi^2(\nu)\right) = \nu = degrees\ of\ freedom$$

$$Var\ \left(\chi^2(\nu)\right) = 2\nu$$

$$Mode\ \left(\chi^2(\nu)\right) = \nu - 2$$

2.3. **The distribution function of $\chi^2(\nu)$.**

$$F(x; \nu) = \int_0^x f(s; \nu) ds$$

is tabulated in most statistics and econometrics texts.

3. THE STUDENT'S t RANDOM VARIABLE

This distribution was published by William Gosset in 1908. His employer, Guinness Breweries, required him to publish under a pseudonym, so he chose "Student."

3.1. **Relationship of Student's t-Distribution to Normal Distribution.** The ratio

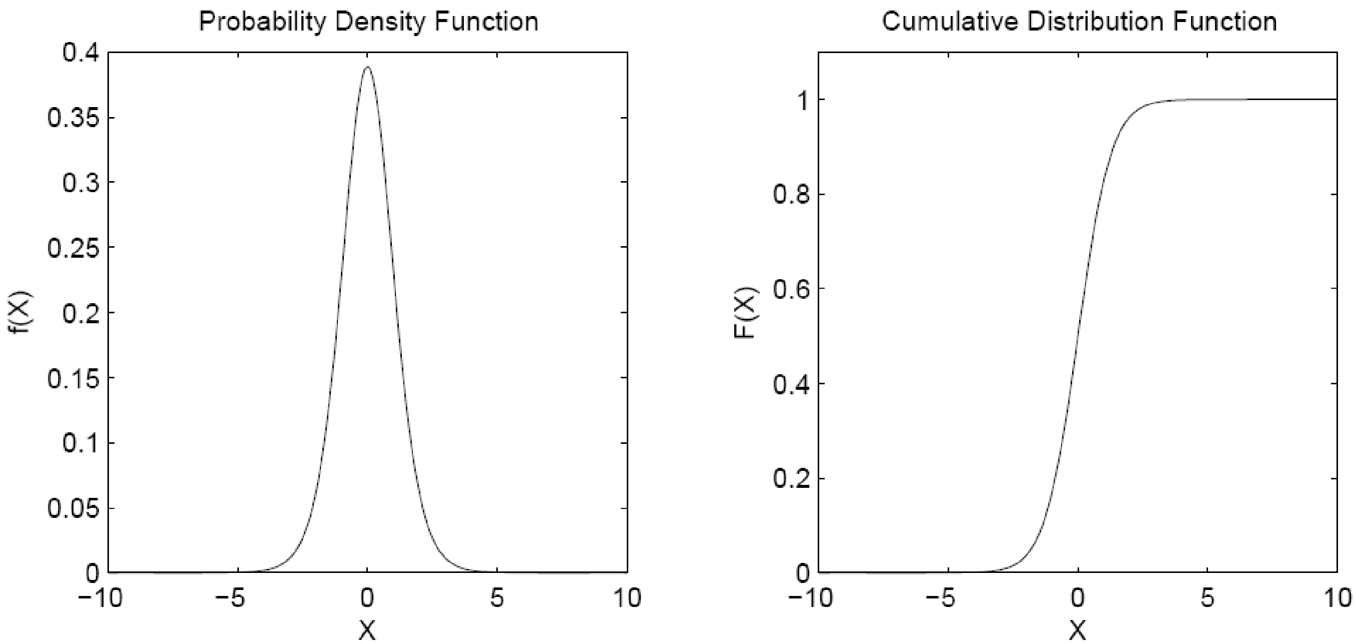
$$t = \frac{N(0,1)}{\sqrt{\frac{\chi^2(\nu)}{\nu}}}$$

has the Student's t density function with ν degrees of freedom where the standard normal variate in the numerator is distributed independently of the χ^{2} variate in the denominator. Tabulations of the associated distribution function are included in most statistics and econometrics books. Note that it is symmetric about origin.

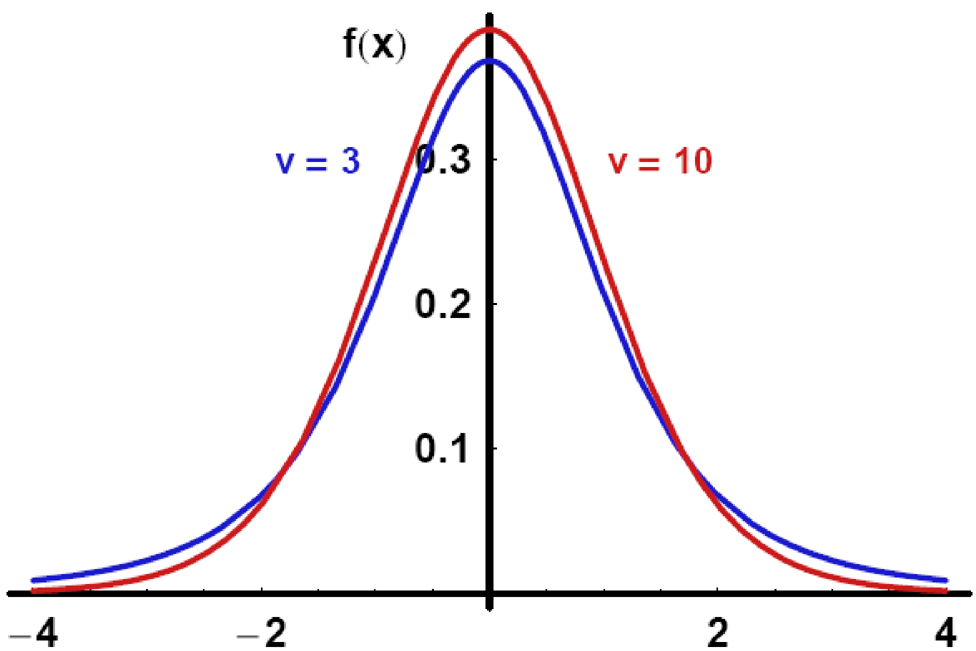
3.2. **Probability Density Function.** The density of Student's t distribution is given by:

$$f(t; \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{\frac{-(\nu+1)}{2}}, \quad -\infty < t < \infty$$

The following diagram shows the pdf and cdf for the Student's t-distribution with parameter $\nu = 10$.



The following diagram shows the cdf for the Student's t-distribution with parameters $\nu = 10$ and $\nu = 3$.



3.3. **Moments of Student's t-distribution.**

$$\begin{aligned} Mean\ (t\,(\nu)) &= 0 \\ Var\ (t\,(\nu)) &= \frac{\nu}{\nu\,-\,2} \end{aligned}$$

4. THE F (FISHER VARIANCE RATIO) STATISTIC

4.1. **Distribution Function.** If $\chi_1^2(\nu_1)$ and $\chi_2^2(\nu_2)$ are independently distributed chi-square variates, then

$$F(\nu_1, \nu_2) = \frac{\frac{\chi_1^2(\nu_1)}{\nu_1}}{\frac{\chi_2^2(\nu_2)}{\nu_2}} = \frac{\nu_2}{\nu_1} \cdot \frac{\chi_1^2(\nu_1)}{\chi_2^2(\nu_2)}$$

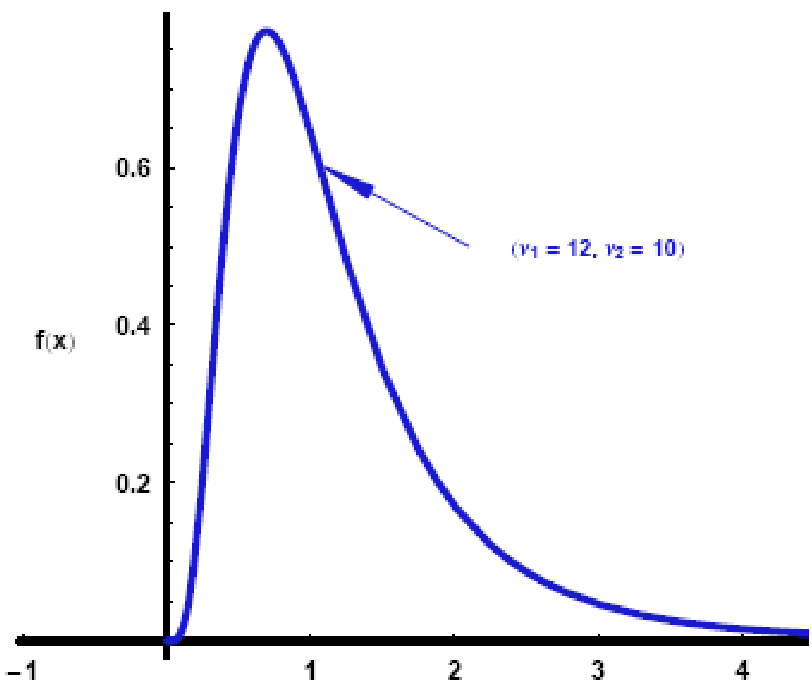
has the F density with ν_1 and ν_2 degrees of freedom.

4.2. **Probability Density Function.** The density of the F distribution is

$$f(F; \nu_1,\, \nu_2) = \begin{cases} \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \cdot \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} \cdot F^{\frac{\nu_1}{2}-1} \cdot \left(1 + \frac{\nu_1}{\nu_2}F\right)^{\frac{-(\nu_1+\nu_2)}{2}} & F\ >\ 0 \\ 0 & otherwise \end{cases}$$

Tabulations of the distribution of $F(\nu_{\,1}, \nu_{\,2})$ are widely available. Note that $F_{\nu_1,\,\nu_2} \sim \left(\frac{1}{F_{\nu_2,\,\nu_1}}\right)$ and therefore the critical values can be found from $f_{\alpha\ \nu_1\,,\ \nu_2} = \left(\frac{1}{f_{1-\alpha\ \nu_2,\ \nu_1}}\right)$.

The first figure below shows the pdf for the F distribution with parameters $\nu_1 = 12$ and $\nu_2 = 20$.



4.3. **Moments of the F distribution.**

$$E(F) = \frac{\nu_2}{\nu_2 - 2}$$

$$Var(F) = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)}$$

5. Exponential Distribution

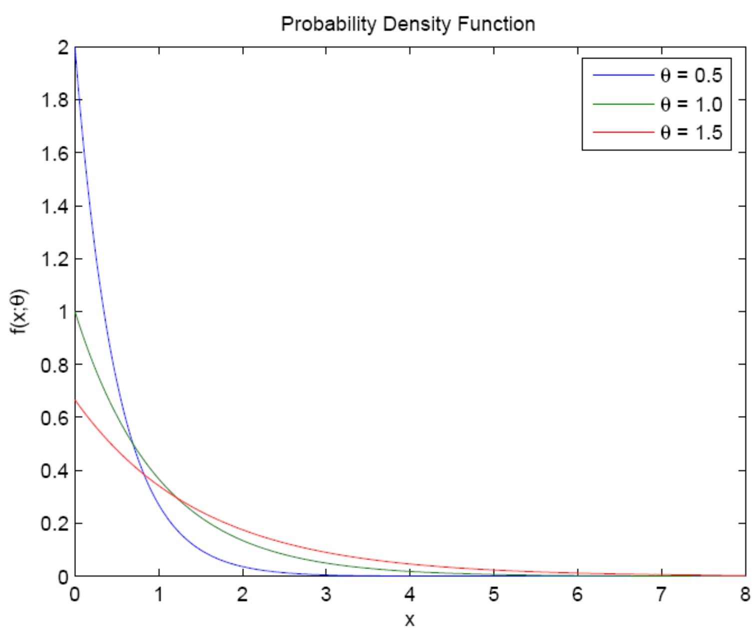
5.1. **Probability density function.** A continuous random variable X has the exponential distribution with parameter $\theta > 0$ if it has a pdf of the form

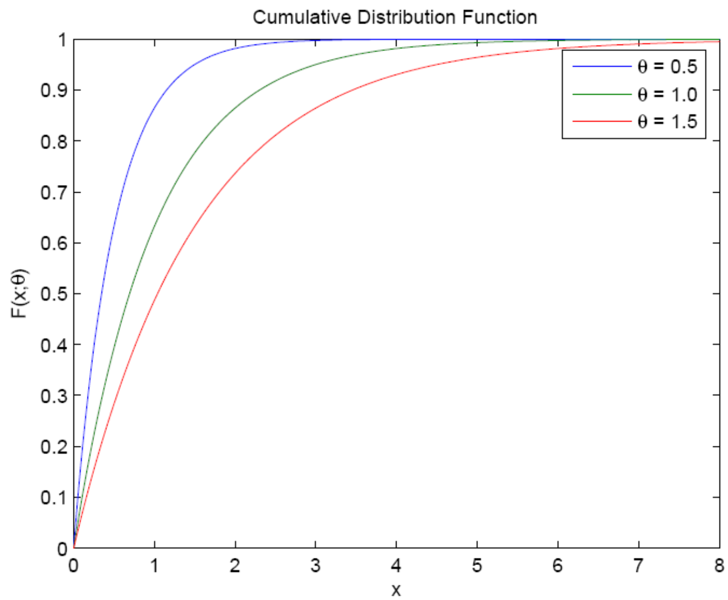
$$\begin{aligned} f(x;\theta) &= \frac{1}{\theta} \cdot e^{-x/\theta} \quad x > 0 \\ &= 0 \end{aligned}$$

The CDF of X is

$$F(x\,;\theta\,) = 1 - e^{-x/\theta} \qquad x > 0$$

The first figure below shows the probability distribution function for the exponential distribution with $\theta = 0.5$, $\theta = 1$, and $\theta = 1.5$ while the second figure shows the cumulative distribution function for the same values of θ.





The exponential distribution, which is an important probability model for lifetime, is sometimes characterized by a property that is given in the following section.

5.2. **Properties of the exponential distribution.** For a continuous random variable X, $X \sim Exp(\theta)$ if and only if

$$P[X > a + t \mid X > a] = P[X > t] \tag{1}$$

for all $a > 0$ and $t > 0$.

Proof (only if)

$$\begin{aligned} P[X &> a + t \mid X > a] = \frac{P[X > a + t \; and \; X > a]}{P[X > a]} \\ &= \frac{P[X > a + t]}{P[X > a]} \\ &= \frac{e^{-(a+t)/\theta}}{e^{-a/\theta}} \\ &= P[X > t] \end{aligned}$$

This shows that the exponential distribution satisfies property (1). This is known as the no-memory property. We will not attempt to show that the exponential distribution is the only such continuous distribution.

If X is the lifetime of a component, then (1) asserts that the probability that the component will last more than $a + t$ time units given that it has already lasted more than a units is the same as that of a new component lasting more than t units. In other words, an old component which is still working is just as reliable as a new component. Failure of such a component is not due to fatigue or wear.

5.2.1. *Moment generating function.* The first moment of the exponential distribution is

$$E\left(X\right)=\int_0^\infty x\frac{1}{\theta}\cdot e^{-x/\theta}dx$$

using the substitution

$$\begin{aligned} u &= \frac{x}{\theta},\ du=\frac{1}{\theta}dx \\ dv &= e^{-x/\theta}dx,\ v=-\theta e^{-x/\theta} \end{aligned}$$

we get

$$\begin{aligned} E\left(X\right) &= \int_0^\infty x\frac{1}{\theta}\cdot e^{-x/\theta}dx \\ &= -xe^{-x/\theta}|_0^\infty+\int_0^\infty e^{-x/\theta}dx \\ &= 0-\theta e^{-x/\theta}|_0^\infty \\ &= \theta \end{aligned}$$

The second raw moment is

$$E\left(X^2\right)=\int_0^\infty x^2\frac{1}{\theta}\cdot e^{-x/\theta}dx$$

using the substitution

$$\begin{aligned} u &= \frac{x^2}{\theta},\ du=2x\frac{1}{\theta}dx \\ dv &= e^{-x/\theta}dx,\ v=-\theta e^{-x/\theta} \end{aligned}$$

we get

$$\begin{aligned} E\left(X\right) &= \int_0^\infty x^2\frac{1}{\theta}\cdot e^{-x/\theta}dx \\ &= -x^2e^{-x/\theta}|_0^\infty+2\int_0^\infty xe^{-x/\theta}dx \\ &= 0+2\theta\int_0^\infty x\frac{1}{\theta}e^{-x/\theta}dx \\ &= 2\theta E\left(X\right)=2\theta^2 \end{aligned}$$

It follows that

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

6. Gamma Distribution

6.1. **Probability density function.** The general formula for the probability density function of the gamma distribution is

$$f(x;\gamma,\beta)=\frac{(\frac{x-\mu}{\beta})^{\gamma-1}e^{(-\frac{x-\mu}{\beta})}}{\beta\Gamma(\gamma)}\qquad x\geqslant\mu,\gamma>0,\beta>0$$

where γ is the shape parameter, μ is the location parameter and β is the scale parameter. Γ is the gamma function which is defined as

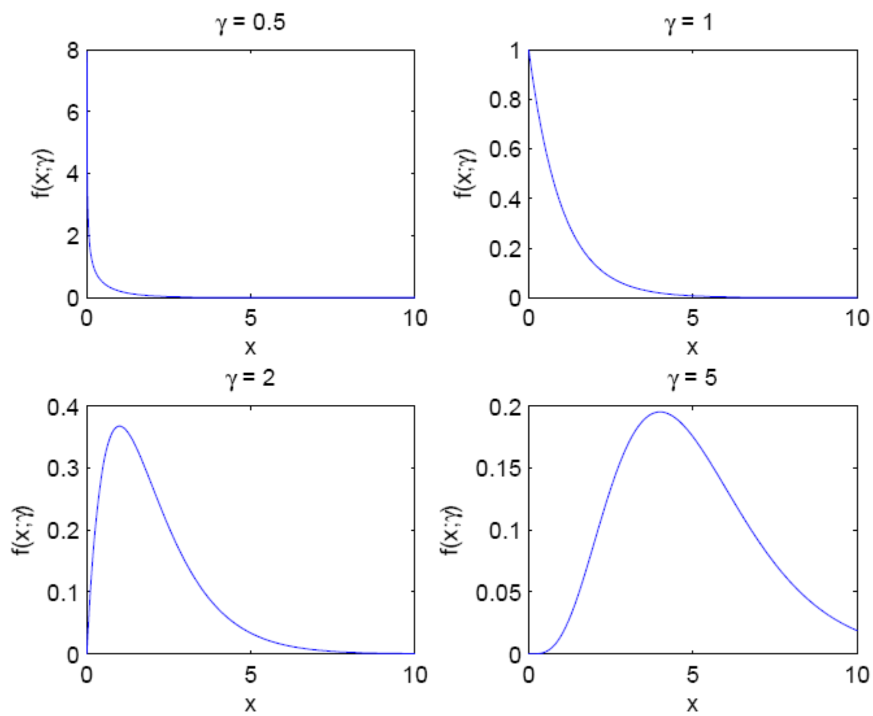
$$\Gamma(a)=\int_0^\infty t^{a-1}e^{-t}dt$$

Most authors only discuss the cases where $\mu = 0$. Therefore, usually we denote the random variable X which has the pdf form above as $X \sim Gamma(\gamma, \beta)$.

When $\mu = 0$ and $\beta = 1$

$$f(x;\gamma) = \frac{(x)^{\gamma-1}e^{(-x)}}{\Gamma(\gamma)}$$

the case is called the standard gamma distribution. The graph of the probability density function is given in the figure below



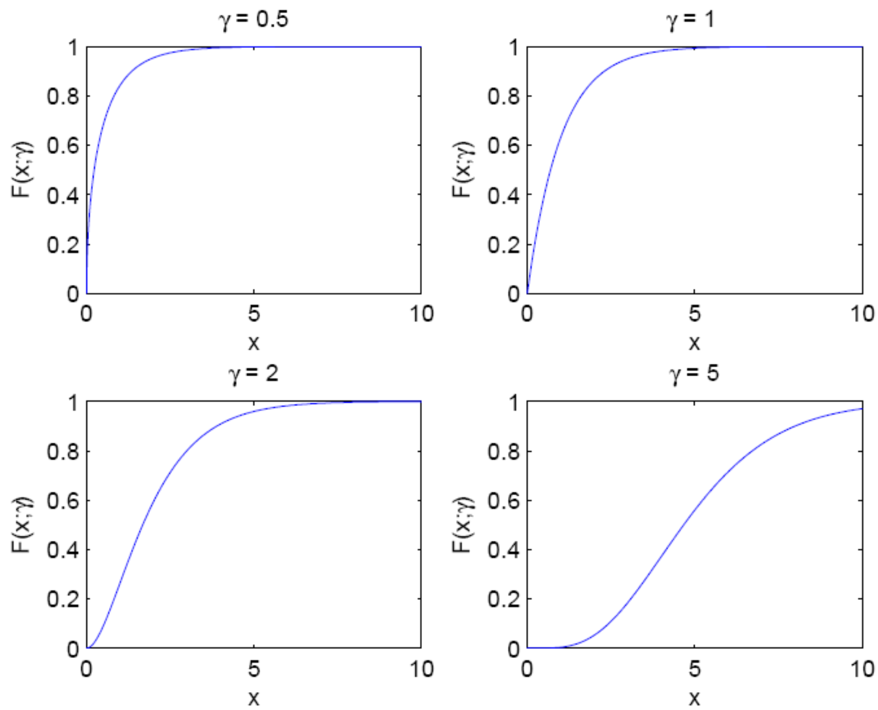
6.2. **Cumulative distribution function.** Correspondingly, the cumulative distribution function (Cdf) of standard gamma distribution is

$$F(x;\gamma) = \frac{\Gamma_x(\gamma)}{\Gamma(\gamma)} \qquad x \geqslant 0, \gamma > 0.$$

where $\Gamma_x(\gamma)$ is the incomplete gamma function, defined as

$$\Gamma_x(\gamma) = \int_0^x t^{\gamma-1}e^{-t}dt$$

The Cdf is plotted with various values of γ in the figure below.



6.3. **Properties of gamma distribution.** When $\mu = 0$, the pdf of gamma distribution is

$$f(x;\gamma,\beta) = \frac{(x)^{\gamma-1}e^{(-\frac{x}{\beta})}}{\beta^{\gamma}\Gamma(\gamma)} \qquad x \geqslant 0, \gamma > 0, \beta > 0$$

which has the following properties.

- If $X_i \sim Gamma(\gamma,\beta)$ for $i = 1, 2, ..., N$, and $\bar{\gamma} = \sum_{i=1}^{N} \gamma_i$, then

$$Y = \sum_{i=1}^{N} X_i \quad \sim Gamma(\bar{\gamma},\beta)$$

  provided that X_i is independently distributed.
- If $X \sim Gamma(\gamma,\beta)$, then $\frac{X}{\beta} \sim Gamma(\gamma, 1)$.
- If $X \sim Gamma(1,\beta)$, then $X \sim Exp(\beta)$, i.e. X is exponentially distributed.
- If $X \sim Gamma(\gamma = \delta/2, \beta = 2)$, then $X \sim \chi^2(\delta)$, X is chi-square distributed.

6.4. **Moments.** The first raw moment is

$$\begin{aligned} E\left(X\right) &= \int_0^\infty x \frac{(x)^{\gamma-1} e^{(-\frac{x}{\beta})}}{\beta^\gamma \Gamma(\gamma)} dx \\ &= \frac{1}{\beta^\gamma \Gamma(\gamma)} \int_0^\infty (x)^\gamma e^{(-\frac{x}{\beta})} dx \quad (let \frac{x}{\beta} = t, x = \beta t,\ dx = \beta dt) \\ &= \frac{1}{\beta^\gamma \Gamma(\gamma)} \int_0^\infty (\beta t)^\gamma e^{(-t)} \beta dt \\ &= \frac{\beta^{\gamma+1}}{\beta^\gamma \Gamma(\gamma)} \int_0^\infty (t)^{(\gamma+1)-1} e^{(-t)} dt \end{aligned}$$

Recall that $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$, so

$$\begin{aligned} E\left(X\right) &= \frac{\beta^{\gamma+1}}{\beta^\gamma \Gamma(\gamma)} \int_0^\infty (t)^{(\gamma+1)-1} e^{(-t)} dt \\ &= \frac{\beta^{\gamma+1} \Gamma(\gamma+1)}{\beta^\gamma \Gamma(\gamma)} \\ &= \frac{\gamma \beta^{\gamma+1} \Gamma(\gamma)}{\beta^\gamma \Gamma(\gamma)} \\ &= \gamma\beta \end{aligned}$$

The second raw moment is

$$E\left(X^2\right) = (\gamma+1)\beta^2\gamma$$

It follows that

$$Var[X] = E\left(X^2\right) - E^2\left(X\right) = (\gamma+1)\beta^2\gamma - (\beta\gamma)^2 = \gamma\beta^2$$