SOME SPECIFIC PROBABILITY DISTRIBUTIONS

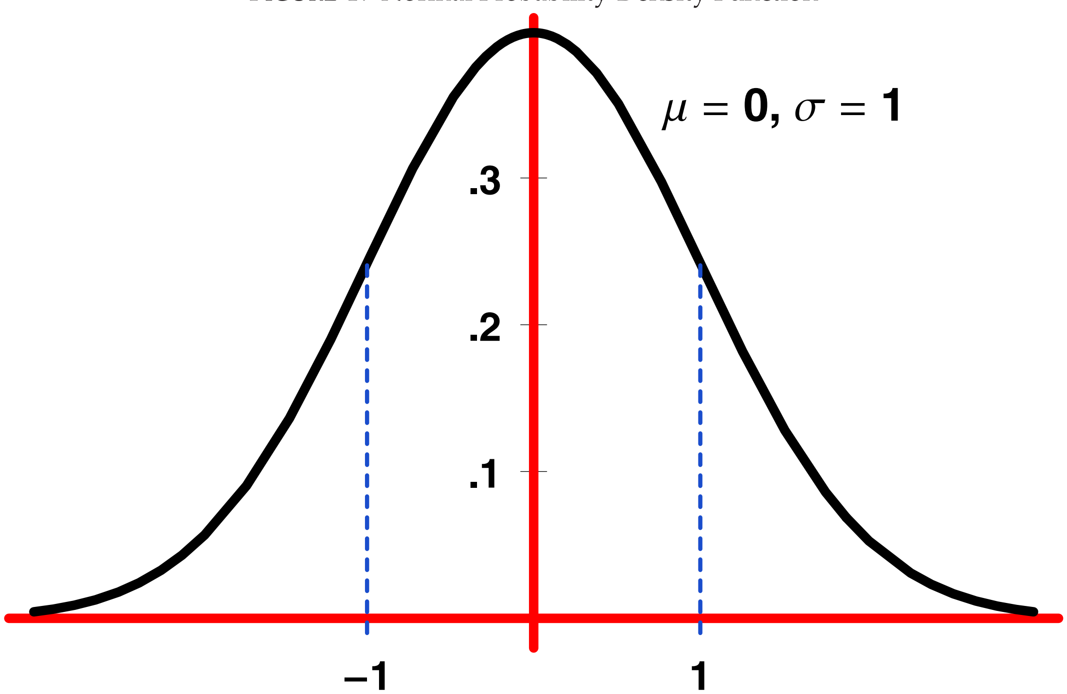
1. Normal random variables

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[μ, σ^2] if the density function of x is given by

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

The normal probability density function is bell-shaped and symmetric. The figure below shows the probability distribution function for the normal distribution with a μ = 0 and σ =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.

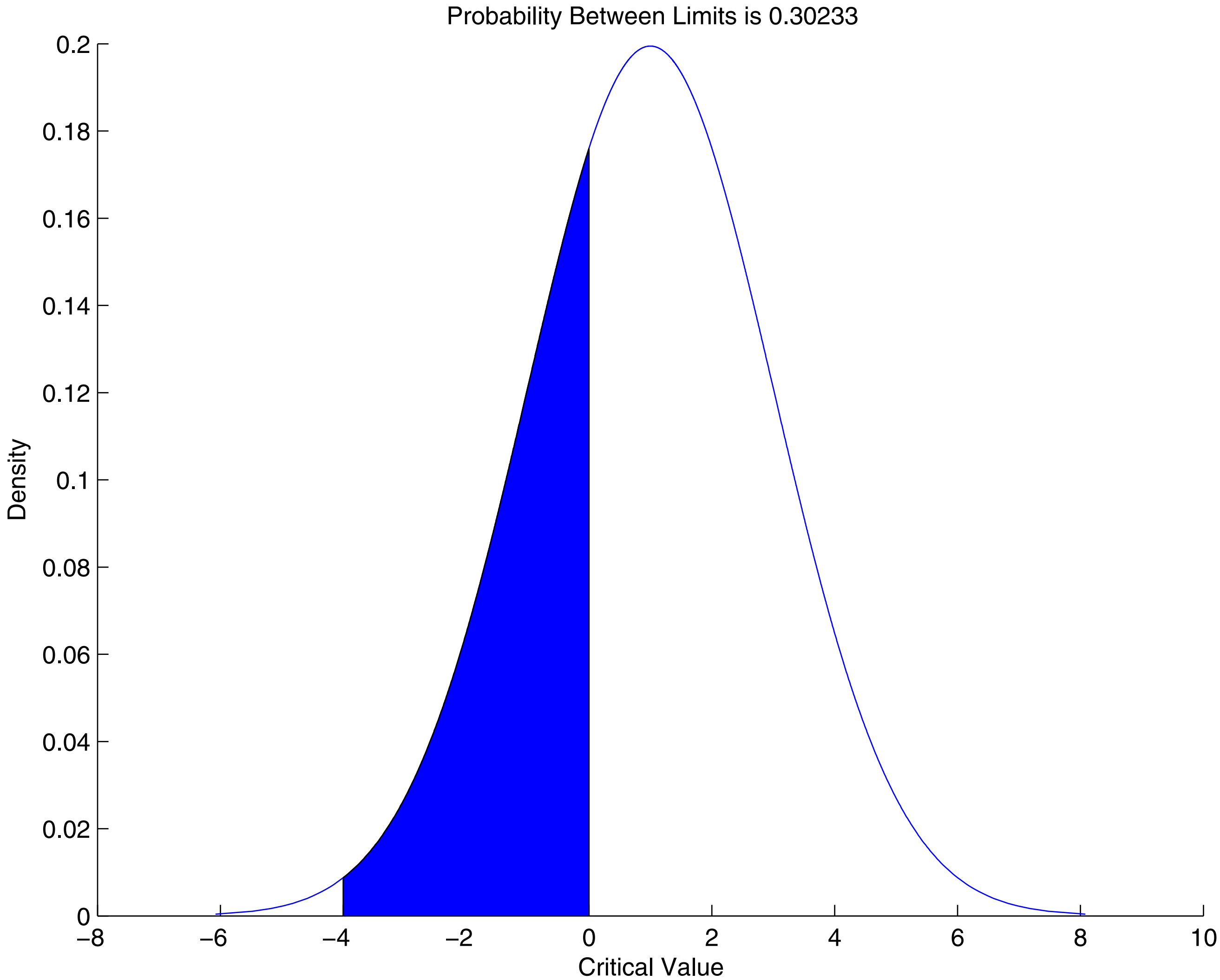
Figure 1. Normal Probability Density Function



Date: November 1, 2005.

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

FIGURE 2. Normal Probability Density Function Showing P$(-4 < x < 0)$



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

a: E(x) = μ, Var(x) = σ^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 σ^2.

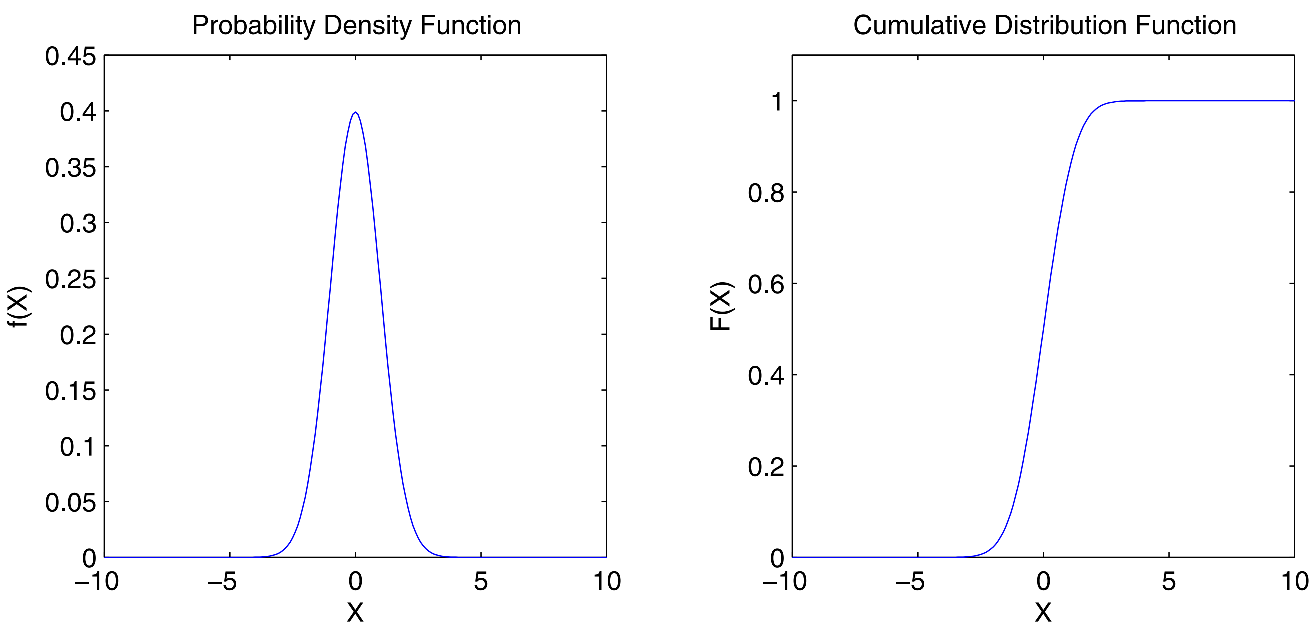
g: The sum of two independently distributed normal random variables is normally distributed. If Y = $\alpha X_1 + \beta X_2 + \gamma$ where $X_1 \sim N(\mu_1, {\sigma_1}^2)$ and $X_2 \sim N(\mu_2, {\sigma^2}_2)$ and X_1 and X_2 are independent, then $Y \sim N(\alpha\mu_1 + \beta\mu_2 + \gamma;\ \alpha^2{\sigma^2}_1 + \beta^2{\sigma^2}_2)$.

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

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

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

FIGURE 3. Normal pdf and cdf



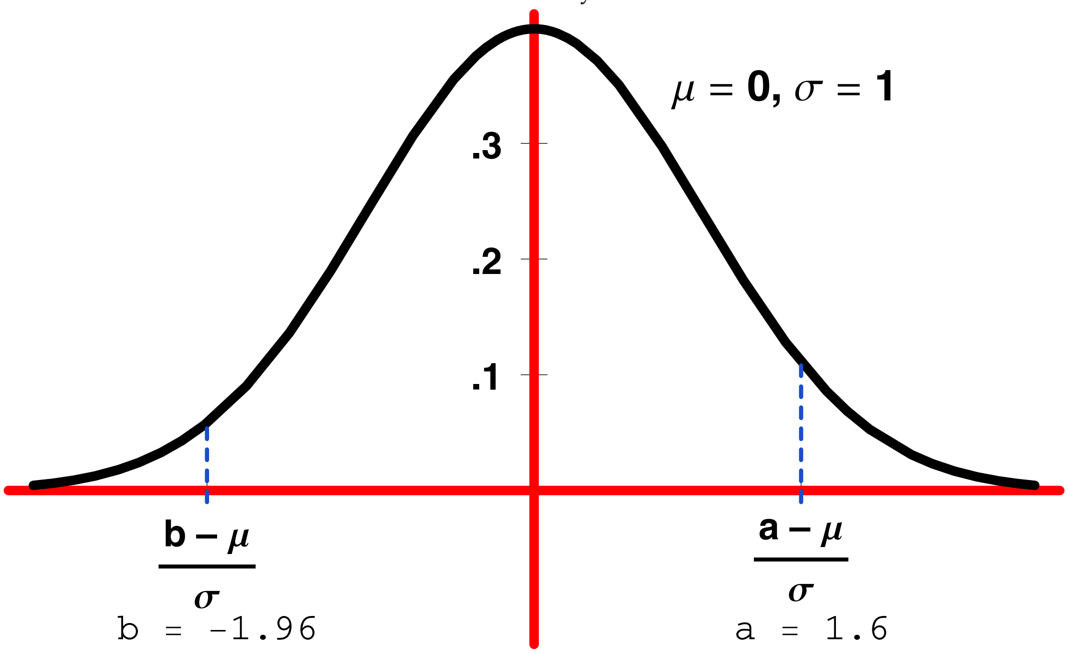
1.4. **Evaluating probability statements with a normal random variable.** If $x \sim N(\mu,\sigma^2)$ then,

$$\begin{aligned} Z\; &= \tfrac{X-\mu}{\sigma} \;\sim\; N(0,1) \\ E\,(Z)\; &= E\left(\tfrac{X-\mu}{\sigma}\right) \;=\; \tfrac{1}{\sigma}\cdot(E(X)-\mu) \;=\; 0 \\ Var\,(Z)\; &= Var\left(\tfrac{X-\mu}{\sigma}\right) \;=\; \tfrac{1}{\sigma^2}Var(X-\sigma) \\ &= \tfrac{\sigma^2}{\sigma^2} \;=\; 1 \end{aligned} \tag{3}$$

Consequently,

$$\begin{aligned} Pr(a \leq x \leq b) &= Pr\,(a - \mu \leq x - \mu \leq b - \mu) \\ &= Pr\left[\frac{a-\mu}{\sigma} \leq \frac{x-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right] \\ &= F\left(\frac{b-\mu}{\sigma};\ 0,\ 1\right) - F\ \left(\frac{a-\mu}{\sigma};\ 0,\ 1\right) \\ &= area\ below \end{aligned} \tag{4}$$

FIGURE 4. Probability of Intervals



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

1.5. **Moment generating function of a normal random variable.** The moment generating function for the central moments is as follows

$$M_X(t) = e^{\frac{t^2\,\sigma^2}{2}}. \tag{5}$$

The first central moment is

$$\begin{aligned} E\,(X - \mu) &= \frac{d}{dt}\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\,|_{t=0} \\ &= t\,\sigma^2\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\,|_{t=0} \\ &= 0 \end{aligned} \tag{6}$$

The second central moment is

$$
\begin{aligned}
E\,(X\,-\,\mu\,)^2 &= \tfrac{d^2}{dt^2}\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\,|_{t\,=\,0} \\
&= \tfrac{d}{dt}\,\left(t\,\sigma^2\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\right)\,|_{t\,=\,0} \\
&= \left(t^2\,\sigma^4\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\,+\,\sigma^2\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\right)\,|_{t\,=\,0} \\
&= \sigma^2
\end{aligned}
\tag{7}
$$

The third central moment is

$$
\begin{aligned}
E\,(X\,-\,\mu\,)^3 &= \tfrac{d^3}{dt^3}\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\,|_{t\,=\,0} \\
&= \tfrac{d}{dt}\,\left(t^2\,\sigma^4\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\,+\,\sigma^2\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\right)\,|_{t\,=\,0} \\
&= \left(t^3\,\sigma^6\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\,+\,2\,t\,\sigma^4\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\,+\,t\,\sigma^4\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\right)\,|_{t\,=\,0} \\
&= \left(t^3\,\sigma^6\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\,+\,3\,t\,\sigma^4\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\right)\,|_{t\,=\,0} \\
&= 0
\end{aligned}
\tag{8}
$$

The fourth central moment is

$$
\begin{aligned}
E\,(X\,-\,\mu\,)^4 &= \tfrac{d^4}{dt^4}\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\,|_{t\,=\,0} \\
&= \tfrac{d}{dt}\,\left(t^3\,\sigma^6\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\,+\,3\,t\,\sigma^4\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\right)\,|_{t\,=\,0} \\
&= \left(t^4\,\sigma^8\,\left(e^{\frac{t^2\sigma^2}{2}}\right)\,+\,3\,t^2\,\sigma^6\,\left(e^{\frac{t^2\sigma^2}{2}}\right)\,+\,3\,t^2\,\sigma^6\,\left(e^{\frac{t^2\sigma^2}{2}}\right)\,+\,3\,\sigma^4\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\right)\,|_{t\,=\,0} \\
&= \left(t^4\,\sigma^8\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\,+\,6\,t^2\,\sigma^6\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\,+\,3\,\sigma^4\,\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\right)\,|_{t\,=\,0} \\
&= 3\,\sigma^4
\end{aligned}
\tag{9}
$$

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

$$
\begin{aligned}
f\;(x\,;\;\nu) &= \tfrac{1}{2^{\frac{\nu}{2}}\,\Gamma\left(\frac{\nu}{2}\right)}\,x^{\frac{\nu-2}{2}}\,e^{\frac{-x}{2}}\;0\;<\;x \\
&= 0\;otherwise
\end{aligned}
\tag{10}
$$

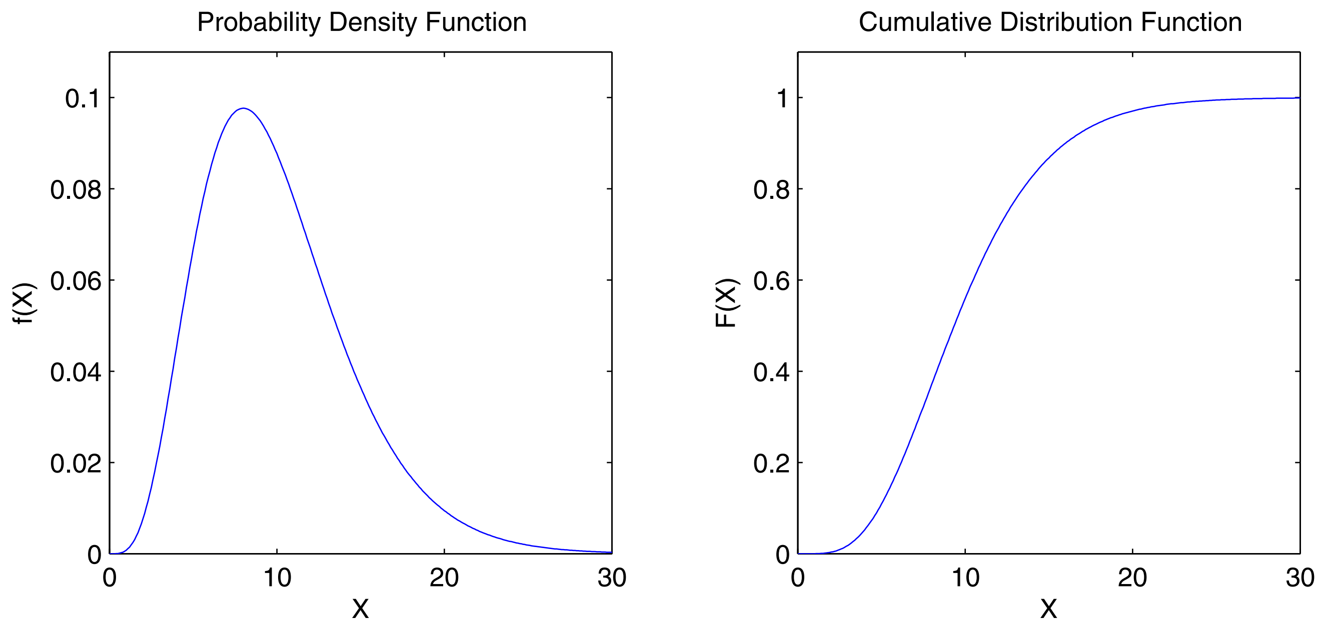
where $\Gamma\,(\,\cdot\,)$ is the gamma function defined by

$$
\Gamma\;(r\,)\;=\int_0^\infty\;u^{\,r\,-\,1}\;e^{\,-u}\;du\,r\;>\;0 \tag{11}
$$

Note that for positive integer values of r, Γ(r) = (r - 1)!

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

FIGURE 5. Chi-square pdf and cdf



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

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

$$\begin{array}{c} If\ X_i \sim N\,(0,\,1)\; i \;=\; 1,\,2,\,\ldots,\,n \\ then\ \sum_{i=1}^{n}\; X_i^2 \;\sim\; \chi^2(n) \end{array} \tag{12}$$

It can also be shown that

$$\begin{array}{c} If\ X_i \sim N\,(0,\,1)\; i \;=\; 1,\,2,\,\ldots,\,n \\ then\ \sum_{i=1}^{n}\; (X_i - \bar{X})^2 \;\sim\; \chi^2(n-1) \end{array} \tag{13}$$

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 *and N(μ,σ^2).*

$$\begin{array}{c} If\ X_i \;\sim\; N\,(\mu,\;\sigma^2)\; i \;=\; 1,\,2,\,\ldots,\,n \\ then\ \displaystyle\sum_{i=1}^{n}\left(\frac{X_i -\;\mu}{\sigma}\right)^2 \;\sim\; \chi^2(n) \\ and\ \displaystyle\sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{\sigma}\right)^2 \;\sim\; \chi^2\,(n-1) \end{array} \tag{14}$$

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

$$y_1 \; + \; y_2 \; \sim \; \chi^2(\nu 1 + \nu 2). \tag{15}$$

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

$$\begin{aligned} Mean\,(\chi^2\,(\nu)) \; &= \; \nu \; = \; degrees\, of\, freedom \\ Var\,(\chi^2\,(\nu)) \; &= \; 2\,\nu \\ Mode\,(\chi^2\,(\nu)) \; &= \; \nu \; - \; 2 \end{aligned} \tag{16}$$

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

$$F(x;\nu) \; = \; \int_0^x \; f\;(s;\nu)ds \tag{17}$$

is tabulated in most statistics and econometrics texts.

2.4. **Moment generating function.** The moment generating function is as follows

$$M_X(t) \; = \frac{1}{(1 \; - \; 2\,t)^{\; \upsilon/2}}\,, t \; < \; \frac{1}{2} \tag{18}$$

The first moment is

$$\begin{aligned} E\,(\,X\,) \; &= \tfrac{d}{dt}\left(\tfrac{1}{(\,1\,-\,2\,t\,)^{\;\upsilon/2}}\right)\;|_{t\,=\,0} \\ &= \left(\tfrac{\upsilon}{(\,1\,-\,2\,t\,)^{\;(\,\upsilon\;+\;1\,)/2}}\right)\;|_{t\,=\,0} \\ &= \upsilon \end{aligned} \tag{19}$$

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

$$t \; = \; \frac{N(0,1)}{\sqrt{\frac{\chi^2(\nu)}{\nu}}} \tag{20}$$

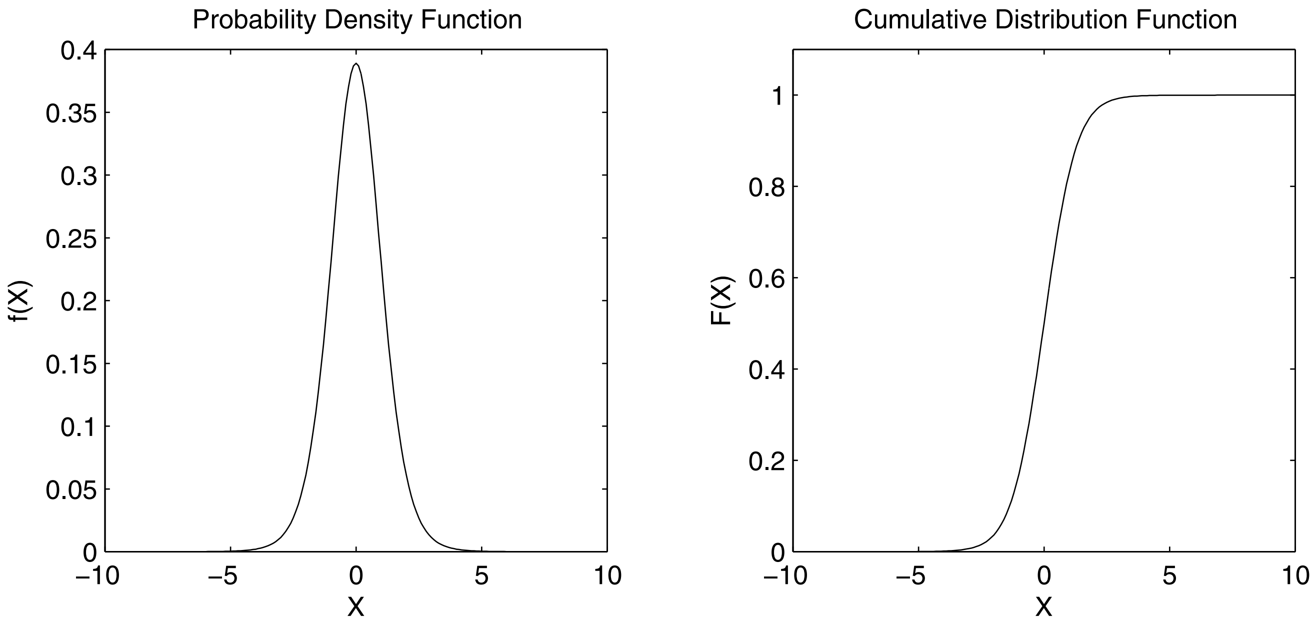
has the Student's t density function with ν degrees of freedom where the standard normal variate in the numerator is distributed independently of the χ^2 variate in the denominator. Tabulations of the associated distribution function are included in most statistics and econometrics books. Note that it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

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

$$f\,(\,t\,;\;\nu\,) \; = \; \frac{\Gamma\left(\frac{\nu\;+\;1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1 + \frac{t^2}{\nu}\right)^{\frac{-(\,\nu\,+\,1\,)}{2}} \quad -\infty \; < \; t \; < \; \infty \tag{21}$$

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

FIGURE 6. Student's t distribution pdf and cdf



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

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

$$\begin{aligned} Mean\ (t\,(\nu)) &= 0 \\ Var\ (t\,(\nu)) &= \frac{\nu}{\nu - 2} \end{aligned} \tag{22}$$

4. THE F (FISHER VARIANCE RATIO) STATISTIC

4.1. **Distribution Function.** If $\chi^2{}_1(\nu_1)$ and $\chi^2{}_2(\nu_2)$ are independently distributed chi-square variates, then

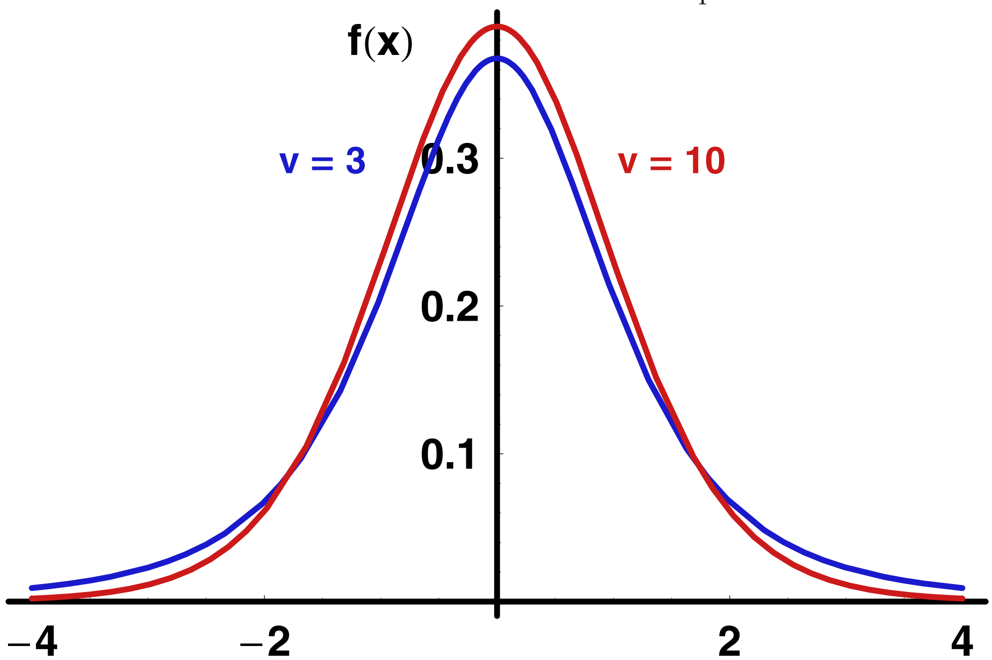
$$F(\nu_1, \nu_2\,) = \frac{\frac{\chi_1^2(\nu_1)}{\nu_1}}{\frac{\chi_2^2(\nu_2)}{\nu_2}} = \frac{\nu_2}{\nu_1} \cdot \frac{\chi_1^2(\nu_1)}{\chi_2^2(\nu_2)} \tag{23}$$

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

$$\begin{aligned} f\,(\,F;\,\nu_1,\,\nu_2) &= \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \cdot \left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}} \cdot F^{\frac{\nu_1}{2}-1} \cdot \left(1 + \frac{\nu_1}{\nu_2}F\right)^{\frac{-(\nu_1+\nu_2)}{2}} \quad F\ >\ 0 \\ &= 0\ otherwise \end{aligned} \tag{24}$$

FIGURE 7. Student's t-distribution with alternative parameter levels



Tabulations of the distribution of F(ν_1,ν_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 following diagram shows the pdf and cdf for the F distribution with parameters ν_1 = 12 and ν_2 = 20.

Here is the pdf of the F distribution for some alternative values of pairs of values (ν_1 and ν_2).

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

$$E(F) = \frac{\nu_2}{\nu_2 - 2} \tag{25}$$

$$Var(F) = \frac{2\nu_2^2(\nu 1 + \nu 2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)} \tag{26}$$

FIGURE 8. Probability of Intervals