PLSC 502 – Autumn 2016
Statistical Distributions II

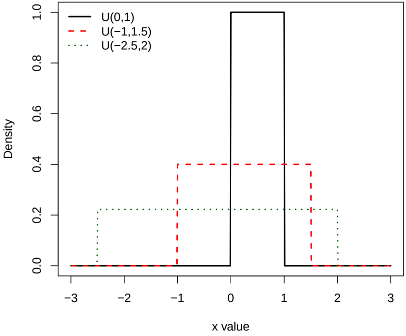
September 29, 2016

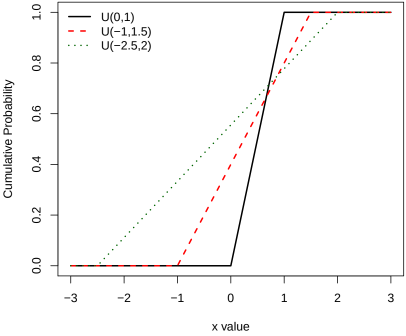
The Uniform Distribution

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b, \\ \\ 0 & \text{for } x < a \text{ or } x > b. \end{cases}$$

$$F(x) = \int f(x)dx = \begin{cases} 0 & \text{for } x < a \\ \\ \frac{x-a}{b-a} & \text{for } a \leq x < b \\ \\ 1 & \text{for } x \geq b \end{cases}$$

Uniform PDFs and CDFs





Uniform Characteristics

$$\mathrm{E}(X) = \check{X} = \frac{a+b}{2}$$

$$\mathrm{mode}(X) = [a, b]$$

$$\mathrm{Var}(X) = \frac{(b-a)^2}{12}$$

$$\mathrm{Skewness}(X) = 0$$

The Standard Uniform Distribution

$$X \sim U(0,1)$$

$$X \sim 1 - X \sim U(0,1).$$

$$F(x) = x.$$

The Normal Distribution

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

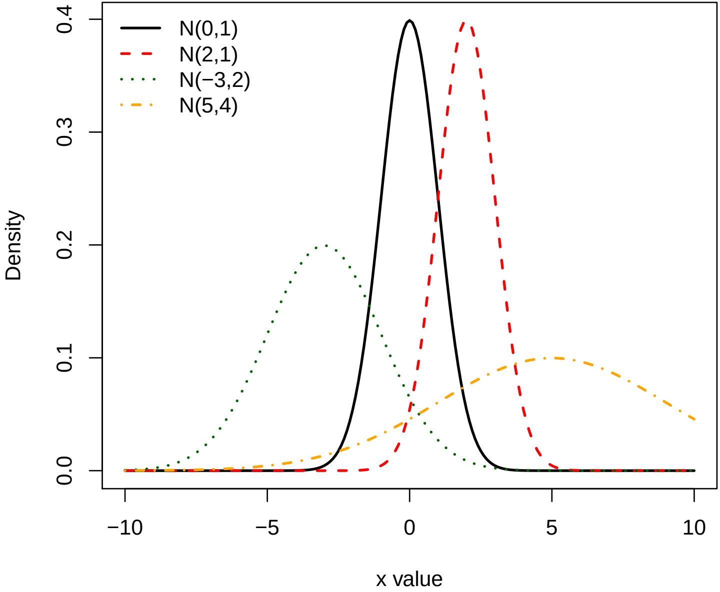
$$X \sim \phi_{\mu,\sigma^2}.$$

$$\begin{aligned} F(x) &= \Phi_{\mu,\sigma^2}(x) \\ &= \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right], \end{aligned}$$

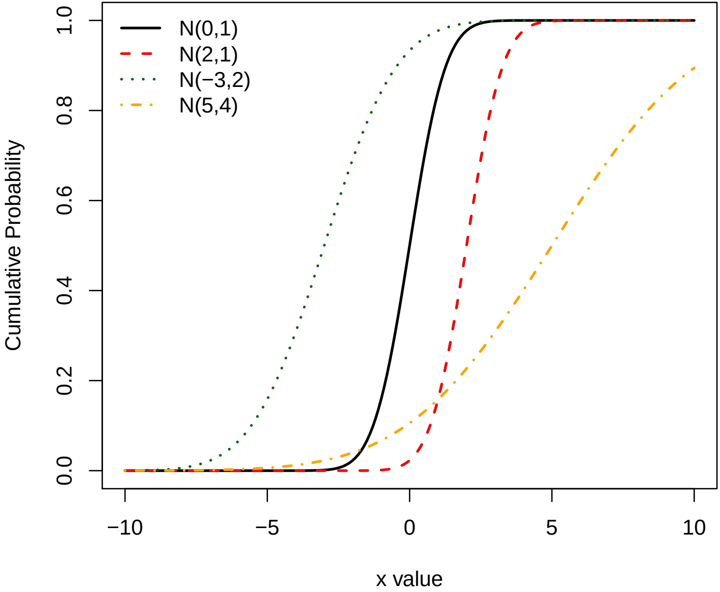
where

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

Various Normal Densities



Various Normal CDFs



Why Normal?

For $i = \{1, 2, ...N\}$ i.i.d. X_i with $\mu_i < \infty$ and $\sigma_i^2 > 0$, define:

$$X = \sum_{i=1}^{N} X_i.$$

Then

$$\begin{aligned} \mathrm{E}(X) &= \sum_{i=1}^{N} \mu_i \\ &= \mu < \infty \end{aligned}$$

and

$$\begin{aligned} \mathrm{Var}(X) &= \sum_{i=1}^{N} \sigma_i^2 \\ &= \sigma^2 < \infty. \end{aligned}$$

Central Limit Theorem

$$\lim_{N \to \infty} X = \lim_{N \to \infty} \sum_{i=1}^{N} X_i \xrightarrow{D} N(\cdot)$$

"...we often think of a normal distribution as being appropriate when the observed variable X can take on a range of continuous values, and when the observed value of X can be thought of as the product of a large number of relatively small, independent "shocks" or perturbations."

Properties of the Normal Distribution

The normal is a two-parameter distribution, where $\mu \in (-\infty, \infty)$ and $\sigma^2 \in (0, \infty)$.

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

- X has support in $\mathfrak{R}$
- $\text{Skewness}(X) = 0$
- X is mesokurtic
- If $X \sim N(\mu, \sigma^2)$, then $aX + b \sim N(a\mu + b, a^2\sigma^2)$

The Standard Normal Distribution

Linear transformation:

- $b = \frac{-\mu}{\sigma}$,
- $a = \frac{1}{\sigma}$.

Yields:

$$\begin{aligned} ax + b &\sim N(a\mu + b, a^2\sigma^2) \\ &\sim N(0,1) \end{aligned}$$

- If $X \sim N(\mu, \sigma^2)$, then $Z = \frac{(x-\mu)}{\sigma} \sim N(0,1)$.
- PDF:

$$f(z) = \equiv \phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(z)^2}{2}\right]$$

Similarly, we often write the CDF for the standard normal as $\Phi(\cdot)$.

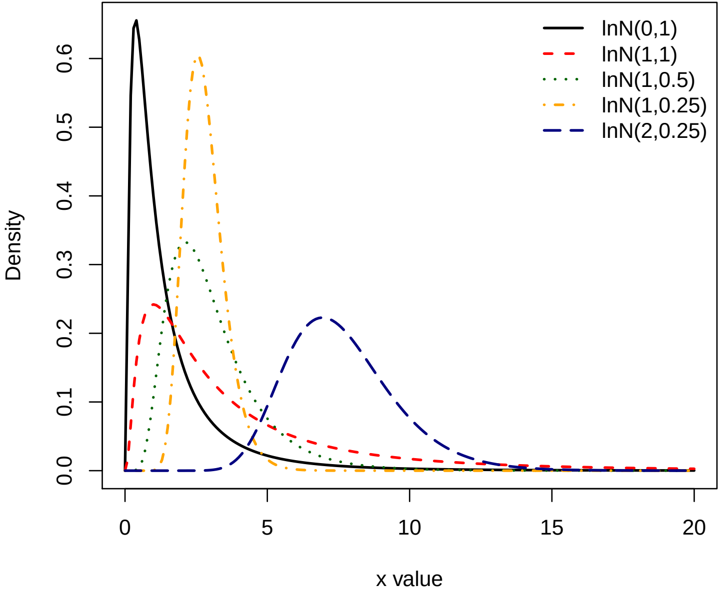
The Log-Normal Distribution

$$Y = \exp(X) \sim \mathrm{LogN}(m, s^2)$$

PDF:

$$f(y) = \frac{1}{ys\sqrt{2\pi}} \times \exp\left[\frac{-(\ln y - m)^2}{2s^2}\right].$$

Log-Normal PDFs



The χ^2 Distribution

PDF:

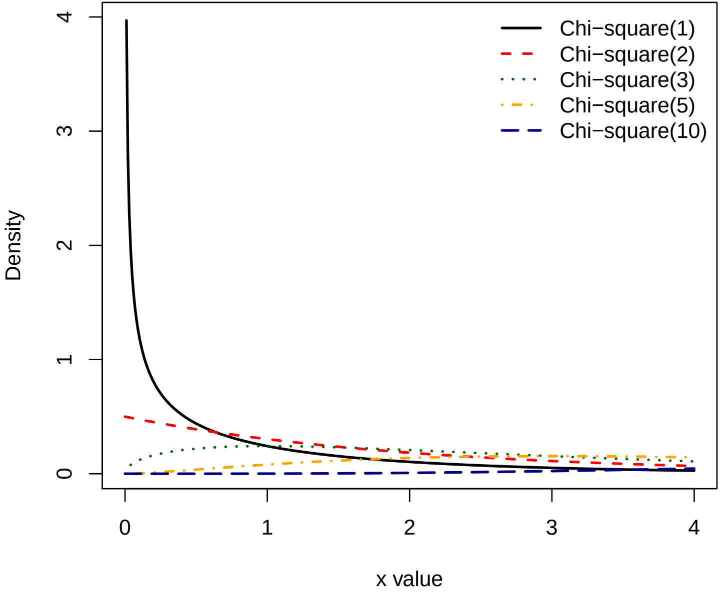
$$\begin{aligned} f(w) &= \frac{1}{2^k \Gamma(k)} w^k \exp\left[\frac{-w}{2}\right] \\ &= \frac{w^{\frac{k-2}{2}} \exp(\frac{-w}{2})}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})} \end{aligned}$$

where $\Gamma(k) = \int_0^\infty t^{k-1} \exp(-t)\, dt$.

CDF:

$$F(w) = \frac{\gamma(k/2, w/2)}{\Gamma(k/2)}$$

χ^2 Densities



χ^2 *Characteristics*

$$E(W) = k$$

$$\text{Var}(W) = 2k.$$

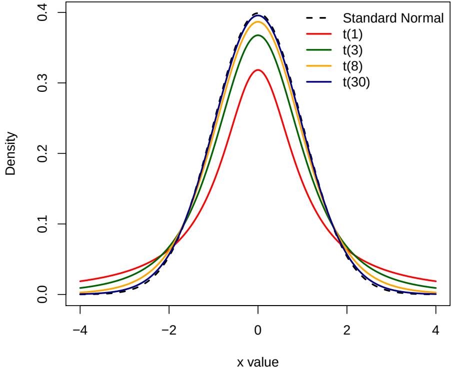
Also:

- If $Z \sim N(0, 1)$, then $Z^2 \sim \chi_1^2$
- If $W_j \sim \chi_j^2$ and $W_k \sim \chi_k^2$ and independent, then
  - $W_j + W_k$ is $\sim \chi_{j+k}^2$ and more generally
  - $\sum_{i=1}^{k} W_i \sim \chi_k^2$.

Student's t Distribution

PDF:

$$f(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{k\pi}\,\Gamma\left(\frac{k}{2}\right)} \left(1 + \frac{x^2}{k}\right)^{-\left(\frac{k+1}{2}\right)}$$



Student's t Characteristics

- $\mu = 0$, $\sigma^2 = \frac{k}{k-2}$
- $t_k \to N(0,1)$ as $k \to \infty$
- If if $Z \sim N(0,1)$, $W \sim \chi^2_k$, and Z and W are independent, then

$$\frac{Z}{\sqrt{W/k}} \sim t_k$$

and

$$\frac{Z^2}{W/k} \sim t_k.$$

The F Distribution

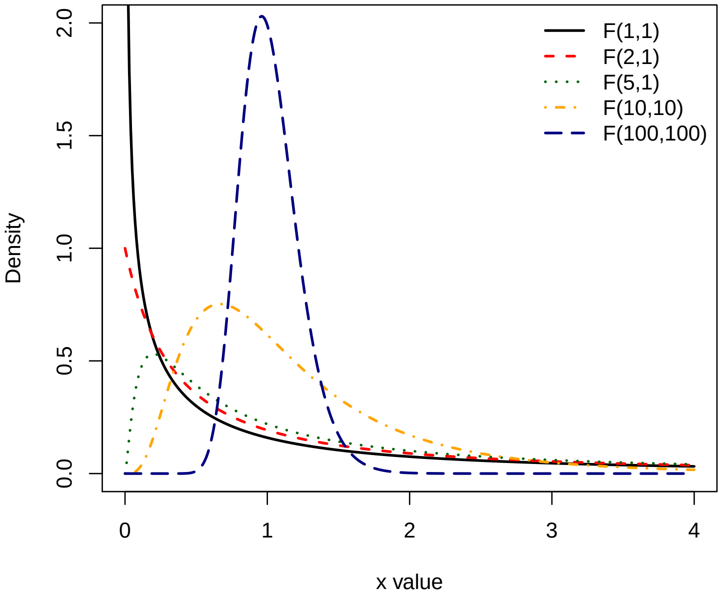
PDF:

$$f(x) = \frac{\left(\frac{k\,x}{k\,x+\ell}\right)^{k/2} \left(1 - \frac{k\,x}{k\,x+\ell}\right)^{\ell/2}}{x\,\mathrm{B}(k/2, \ell/2)}$$

where $\mathrm{B}(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1}\,dt$.

$$X \sim F_{k,\ell}$$

F Densities



F Characteristics

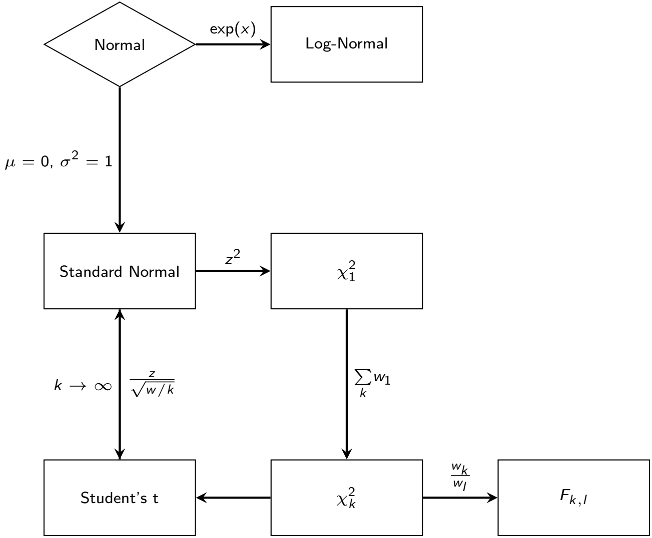
- $\mathrm{E}(X) = \frac{\ell}{\ell-2}$
- $\mathrm{Var}(X) = \frac{2\,\ell^2\,(k+\ell-2)}{k(\ell-2)^2(\ell-4)}$
- For independent $W_1 \sim \chi^2_k$ and $W_2 \sim \chi^2_\ell$:

$$\frac{W_1}{W_2} \sim F_{k,\ell}$$

Relationships Among Continuous Distributions

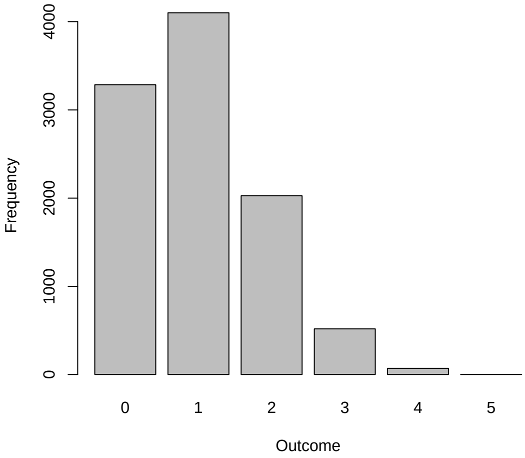
Variables

$$x \sim N(\mu, \sigma^2)$$
$$z \sim N(0, 1)$$
$$w_k \sim \chi^2_k$$



Drawing From Distributions

```
> Xbinom5point2<-rbinom(10000,5,0.2)
```



R and Stata Commands...

Table: Commands for Generating Random Variates

Distribution	R	Stata
Binomial(n, π)	`rbinom()`	`rndbin`*
Geometric(π)	`rgeom()`	?
Negative Binomial(n, π)	`rnbinom()`	?
Poisson(λ)	`rpois()`	`rndpoi`*
Uniform($0, 1$)	`runif()`	`uniform()`
Normal($0, 1$)	`rnorm()`	`invnorm(uniform())`
Lognormal($0, 1$)	`rlnorm()`	`xlgn`*
Student's $t(k)$	`rt()`	`rndt`*
Chi-Square(k)	`rchisq()`	`rndchi`*
$F(k, \ell)$	`rf`	`rndf`*

Note: Stata commands marked with an asterisk are from Hilbe's `rnd` group of commands. "?"s indicate that I'm not aware of any "canned" way of doing this, though one can always generate them "by hand" using the appropriate PDF function.

Pseudo-Random Numbers and "Seeds"

```
> seed<-3229 # calling "seed" some thing
> set.seed(seed) # setting the system seed
> rt(3,1) # three draws from a t distrib. w/1 d.f.
[1] -0.1113 -0.7306  1.9839
> seed<-1077
> set.seed(seed) # resetting the seed
> rt(3,1) # different values for the draws
[1]    -0.5211    7.9161 -155.3186
> seed<-3229 # original seed
> set.seed(seed)
> rt(3,1) # identical values of the draws
[1] -0.1113 -0.7306  1.9839
```