

PLSC 502 – Autumn 2023

Probability Distributions

October 2, 2023

Why learn about probability distributions?

- Things we study are both *regular* and *random*.
- Inference: Making claims about the world without getting *all* the data...
- Prediction: Getting things right (on average), but also knowing how precisely we do that.
- For more answer(s) to this question, see [this](#).

What do we need to know?

Distributions are characterized by a combination of their *formulas* and their *moments*...

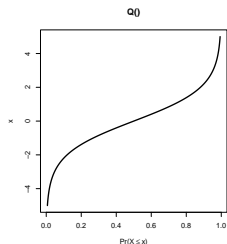
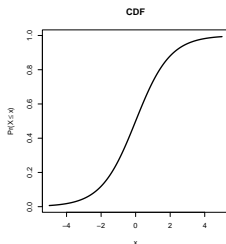
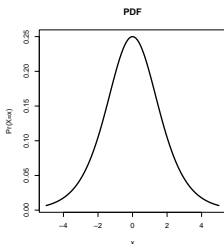
- The *formula* describes the mapping from the outcome space to the probability space.
- The *moments* tell us about the “location,” shape, and other characteristics of a particular instance of a distribution.
 - Most distributions are fully characterized by their formula + one or more (usually a small number of) finite moments (often called “parameters”).
 - A distribution's *moment generating function* provides a bridge between its moments and the formula for the distribution.

Describing Distributions

Attributes of Distributions

Attribute	a/k/a	Meaning	Function	In R
Density	$f(X)$, "PDF" *	$\Pr(X = x)$	Maps from x to $\Pr(X = x)$	<code>ddist</code>
Distribution	$F(X)$, "CDF"	$\Pr(X \leq x) = \int_{-\infty}^x f(X)dX$	Maps from x to $\Pr(X \leq x)$	<code>pdist</code>
Quantile	$Q(p)$, $F_X^{-1}(p)$	Value of x such that $\Pr(X \leq x)$	Maps from $\Pr(X \leq x)$ to x	<code>qdist</code>

* Sometimes called a "probability mass function" ("PMF") when X is discrete.



Discrete Distributions

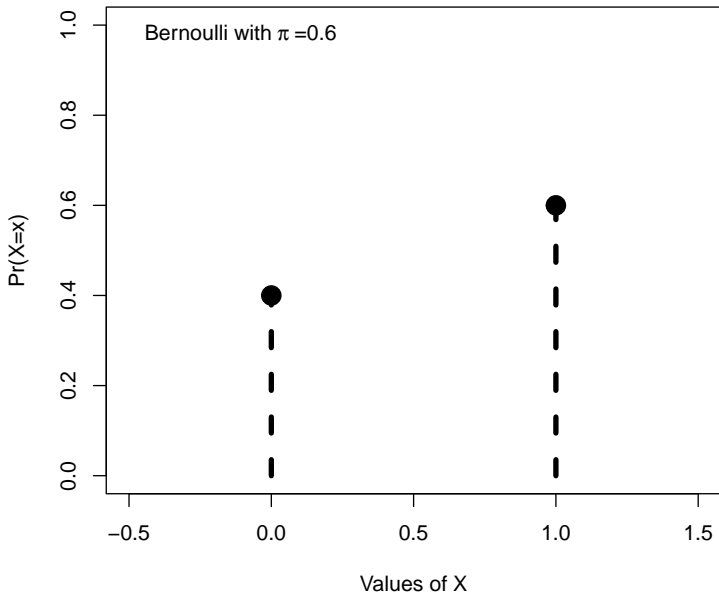
Binary X :

$$\begin{aligned} X &= 0 \text{ with probability } 1 - \pi \\ &= 1 \text{ with probability } \pi. \end{aligned}$$

Density / PDF:

$$\begin{aligned} f(x) &= \begin{cases} 1 - \pi & \text{for } X = 0 \\ \pi & \text{for } X = 1 \end{cases} \\ &= \pi^x (1 - \pi)^{1-x}, \quad x \in \{0, 1\} \end{aligned}$$

$$X \sim \text{Bernoulli}(\pi)$$



CDF:

$$\begin{aligned} F(x) &= \sum_x f(x) \\ &= \begin{cases} 1 - \pi & \text{for } X = 0 \\ 1 & \text{for } X = 1 \end{cases} \end{aligned}$$

Expectation:

$$\begin{aligned} E(X) &= \sum_x xf(x) \\ &= (0)(1 - \pi) + (1)(\pi) \\ &= \pi \end{aligned}$$

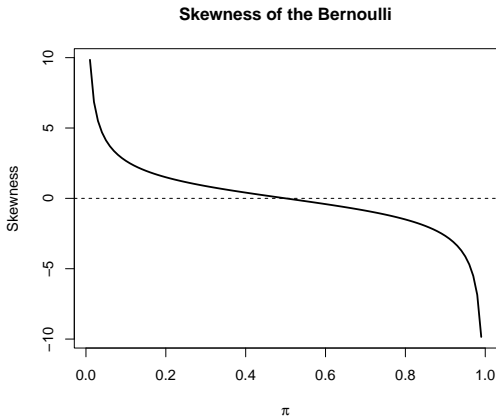
Variance:

$$\begin{aligned}\text{Var}(X) &= \sum_x [X - E(X)]^2 f(x) \\&= \sum_x [X - \pi]^2 f(x) \\&= (0 - \pi)^2(1 - \pi) + (1 - \pi)^2\pi \\&= \pi^2 - \pi^3 + \pi - 2\pi^2 + \pi^3 \\&= \pi - \pi^2 \\&= \pi(1 - \pi)\end{aligned}$$

Even More Bernoulli

Skewness:

$$\text{Skewness} = \frac{(1 - \pi) - \pi}{\sqrt{(1 - \pi)\pi}}$$



Even More Bernoulli

MGF:

$$\begin{aligned}\psi(t) &= \int_{-\infty}^{\infty} \exp(tx) dF(x) \\ &= \sum_{n=0}^1 \exp(tn) \pi^n (1-\pi)^{1-n} \\ &= \exp(0)(1-\pi) + \exp(t)\pi \\ &= (1-\pi) + \pi \exp(t)\end{aligned}$$

Implying:

$$\frac{\partial^k \psi(t)}{\partial^k t} = \pi \exp(t) \quad \forall k$$

and *raw moments*:

$$E(X^k) = \pi \quad \forall k > 0$$

Central moments:

$$M_1 = \pi,$$

$$M_2 = \pi(1 - \pi),$$

$$M_3 = \pi(1 - \pi)(1 - 2\pi),$$

etc.

Assume n independent binary “trials,” each with identical probability of “success” π . Then the number of “successes” in n trials follows a *binomial* distribution:

$$f(x) = \binom{n}{x} \pi^x (1 - \pi)^{n-x}$$

where recall that

$$\binom{n}{x} \equiv \frac{n!}{x!(n-x)!}.$$

$$X \sim \text{binomial}(n, \pi).$$

Why “binomial”?

Binomial Theorem:

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}.$$

e.g.

$$(a + b)^2 = a^2 + 2ab + b^2,$$

$$(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3,$$

$$(a + b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4,$$

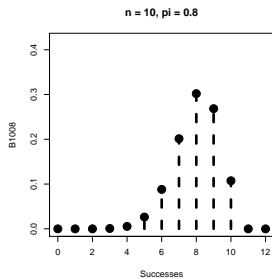
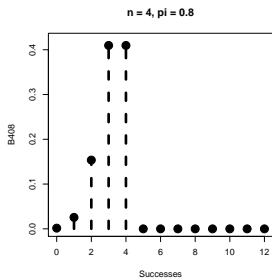
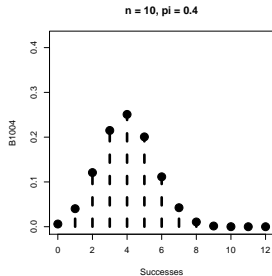
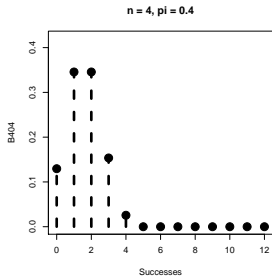
Binomial: Derivation

For $n = 2$:

$$\begin{aligned}\Pr(X = 0) &= \Pr(X_1 = 0, X_2 = 0) \\ &= \Pr(X_1 = 0) \times \Pr(X_2 = 0) \\ &= (1 - \pi)^2\end{aligned}$$

$$\begin{aligned}\Pr(X = 1) &= \Pr(X_1 = 1, X_2 = 0 \text{ or } X_1 = 0, X_2 = 1) \\ &= \Pr(X_1 = 1) \times \Pr(X_2 = 0) + \Pr(X_1 = 0) \times \Pr(X_2 = 1) \\ &= \pi(1 - \pi) + (1 - \pi)\pi \\ &= 2[\pi(1 - \pi)]\end{aligned}$$

$$\begin{aligned}\Pr(X = 2) &= \Pr(X_1 = 1, X_2 = 1) \\ &= \Pr(X_1 = 1) \times \Pr(X_2 = 1) \\ &= \pi^2\end{aligned}$$



CDF:

$$\begin{aligned} F(x) &= \sum_{j=0}^x f(j) \\ &= \sum_{j=0}^x \binom{n}{j} \pi^j (1 - \pi)^{n-j} \end{aligned}$$

Expectation:

$$E(X) = n\pi,$$

Variance:

$$\begin{aligned}\text{Var}(X) &= \sum_x [X - E(X)]^2 f(x) \\ &= \sum_x (X - \pi n)^2 \binom{n}{x} \pi^x (1 - \pi)^{n-x} \\ &= n\pi(1 - \pi).\end{aligned}$$

Skewness:

$$\text{Skewness} = \frac{1 - 2\pi}{\sqrt{n\pi(1 - \pi)}}$$

The Binomial...

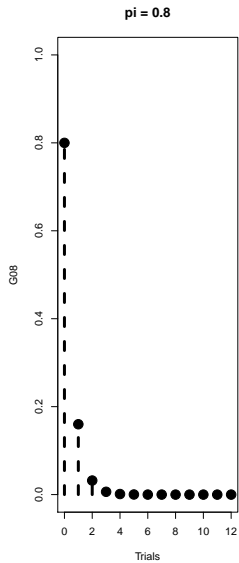
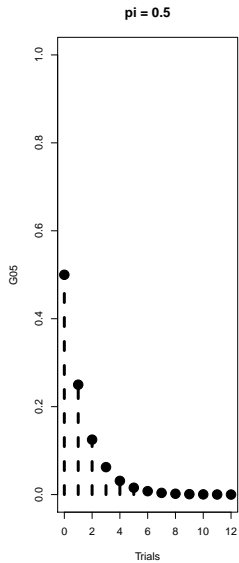
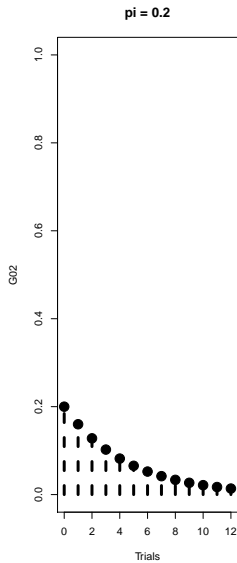
- Is *unimodal* (except in certain cases),
- has median $\lceil n\pi \rceil$ or $\lfloor n\pi \rfloor$,
- has mode $\lceil (n+1)\pi \rceil$ or $\lfloor (n+1)\pi \rfloor$,
- has skewness that is:
 - increasing in n , and
 - is largest when $\pi = 0.5$ for a fixed value of n .

The number of independent Bernoulli trials needed to achieve *one* success is a *geometric* random variable.

PDF:

$$f(x) = \pi(1 - \pi)^{x-1}$$

$$X \sim \text{geometric}(\pi).$$



CDF:

$$\begin{aligned} F(x) &= \sum_{j=1}^x \pi(1-\pi)^{x-1} \\ &= 1 - (1-\pi)^x \end{aligned}$$

Expectation:

$$E(X) = \frac{1}{\pi}$$

Variance:

$$\text{Var}(X) = \frac{1-\pi}{\pi^2}$$

Negative Binomial

The number of *failures we observe* (x) before achieving the r th success in a series of independent Bernoulli trials (each with equal probability of success π) is distributed according to a *negative binomial* distribution.

PDF:

$$f(x) = \binom{r+x-1}{r-1} \pi^r (1-\pi)^x$$

More Negative Binomial

CDF:

$$\begin{aligned} F(x) &= \sum_{j=0}^x \binom{r+j-1}{r-1} \pi^r (1-\pi)^j \\ &= 1 - \text{CDF}_{\text{binomial}} \end{aligned}$$

Expected value:

$$E(X) = \frac{(1-\pi)r}{\pi}$$

Even More Negative Binomial

Variance:

$$\text{Var}(X) = \frac{(1 - \pi)r}{\pi^2}.$$

Skewness:

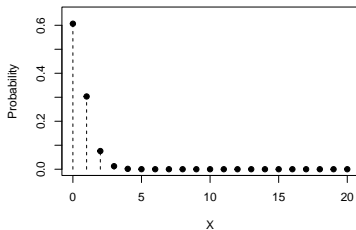
$$\text{Skewness} = \frac{1 + \pi}{\sqrt{\pi r}}$$

For n independent Bernoulli trials with (sufficiently small) probability of success π and where $n\pi \equiv \lambda > 0$, the probability of observing exactly x total “successes” as the number of trials grows without limit is the *Poisson distribution*.

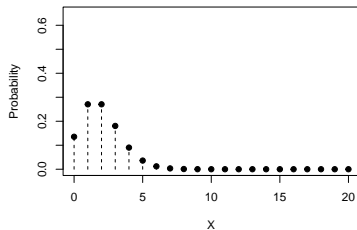
PDF:

$$\begin{aligned} f(x) &= \lim_{n \rightarrow \infty} \left[\binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \right] \\ &= \frac{\lambda^x \exp(-\lambda)}{x!}. \end{aligned}$$

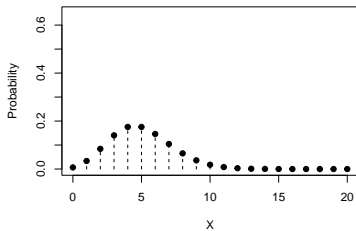
Lambda = 0.5



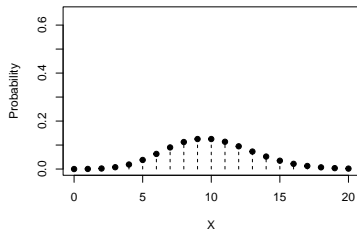
Lambda = 1



Lambda = 4



Lambda = 8



CDF:

$$F(x) = \sum_{j=0}^x \frac{\lambda^j \exp(-\lambda)}{j!}.$$

Mean & Variance:

$$E(X) = \text{Var}(X) = \lambda$$

All higher moments are zero...

Alternative Poisson

Independent, constant-probability events occurring in time...

$$\text{"Arrival rate"} = \lambda$$

Implies:

$$\begin{aligned}\Pr(\text{Event in } (t, t+h]) &= \lambda h \\ \Pr(\text{No event in } (t, t+h]) &= 1 - \lambda h\end{aligned}$$

$$N_{\text{Events occurring in } (t, t+h]} = \frac{\exp(-\lambda h) \lambda h^x}{x!}$$

If $h = 1 \forall h$, then:

$$f(x) = \frac{\exp(-\lambda) \lambda^x}{x!}$$

Multinomial

Imagine K possible distinct *outcomes* for each “trial,” where each possible outcome $k \in \{1, 2, \dots, K\}$ has π_k and $\sum_{k=1}^K \pi_k = 1$.

Define x_k = number of times we observe outcome k out of n trials.

Then for:

$$\mathbf{x} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_K \end{pmatrix}$$

the PDF is:

$$f(\mathbf{x}) = \frac{n!}{x_1! x_2! \dots x_K!} \pi_1^{x_1} \pi_2^{x_2} \dots \pi_K^{x_K}$$

Multinomial, continued

Expected value:

$$E(\mathbf{X}) \equiv E \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_K \end{pmatrix} = n \begin{pmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_K \end{pmatrix}$$

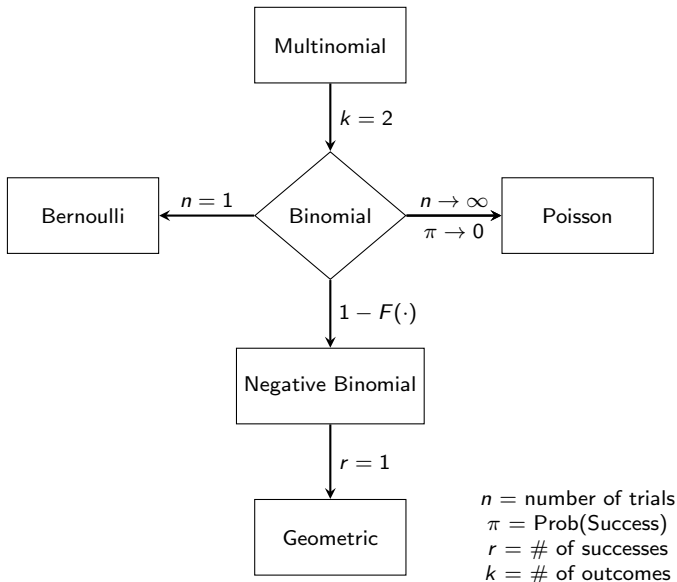
Variance:

$$\text{Var}(\mathbf{X}) \equiv \text{Var} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_K \end{pmatrix} = n \begin{bmatrix} \pi_1(1 - \pi_1) \\ \pi_2(1 - \pi_2) \\ \vdots \\ \pi_K(1 - \pi_K) \end{bmatrix}$$

Covariance between X_s and X_t , $s \neq t$:

$$\text{Cov}(X_s, X_t) = -n\pi_s\pi_t$$

Schematic



Continuous Distributions

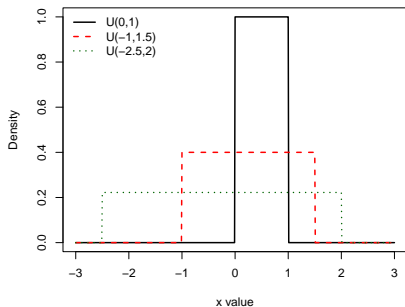
The Uniform Distribution

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

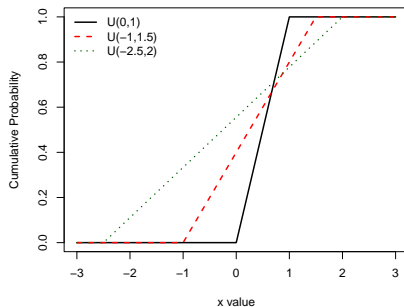
$$F(x) = \int f(x)dx = \begin{cases} 0 & \text{for } x < a \\ \frac{x-a}{b-a} & \text{for } a \leq x < b \\ 1 & \text{for } x \geq b \end{cases}$$

Uniform PDFs and CDFs

Probability Density Functions



Cumulative Distribution Functions



Uniform Characteristics

Expected value / “mean”:

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

Mode:

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

Variance:

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

Skewness:

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

The Standard Uniform Distribution

Special case of the uniform:

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

It has the property that:

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

In addition, the CDF is also unique:

$$F(x) = x$$

The Normal Distribution

The Normal density is:

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

abbreviated:

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

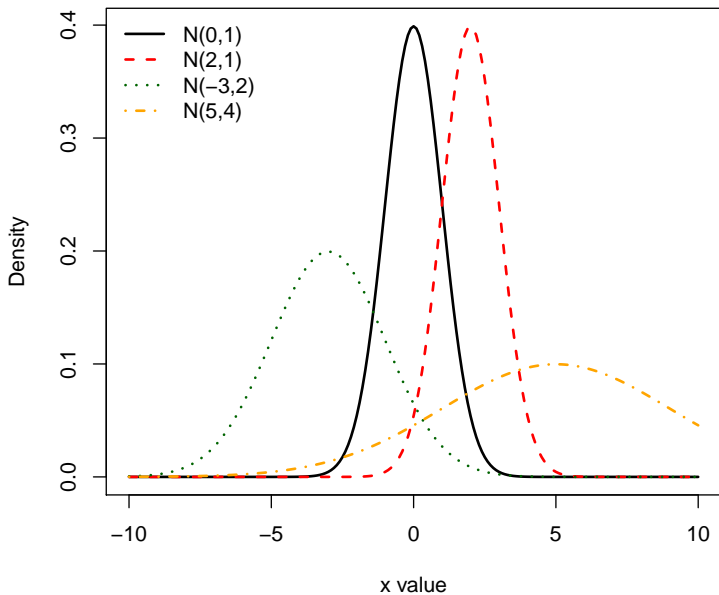
The CDF is:

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

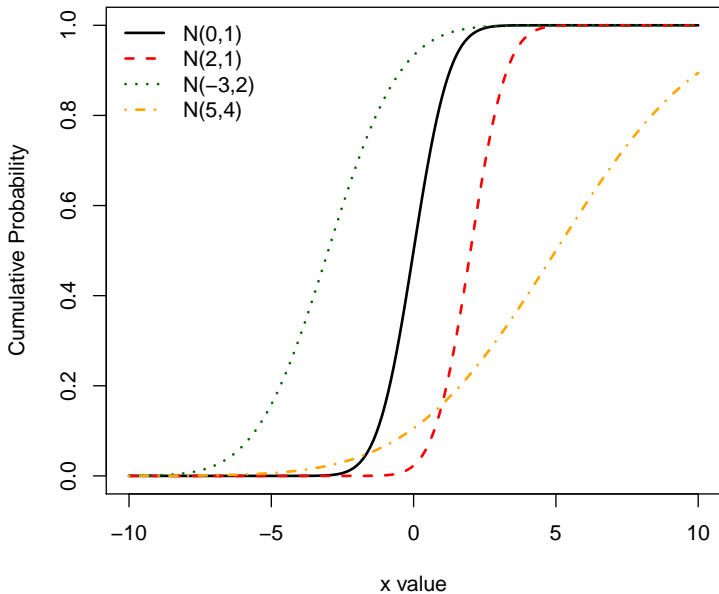
where

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

Various Normal Densities



Various Normal CDFs



Why Normal?

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

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

Then

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

and

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

Central Limit Theorem

The “CLT”:

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

“...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 \Re
- $\text{Skewness}(X) = 0$
- 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:

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

In other words:

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

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

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

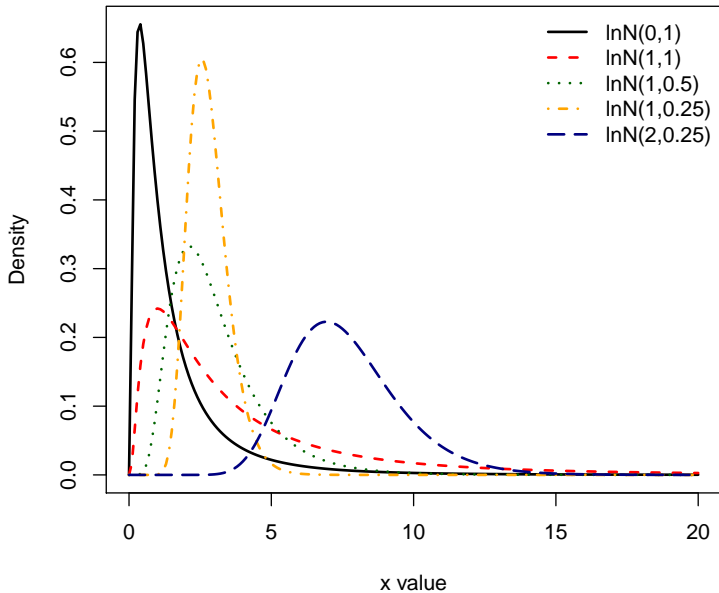
The Log-Normal Distribution

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

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:

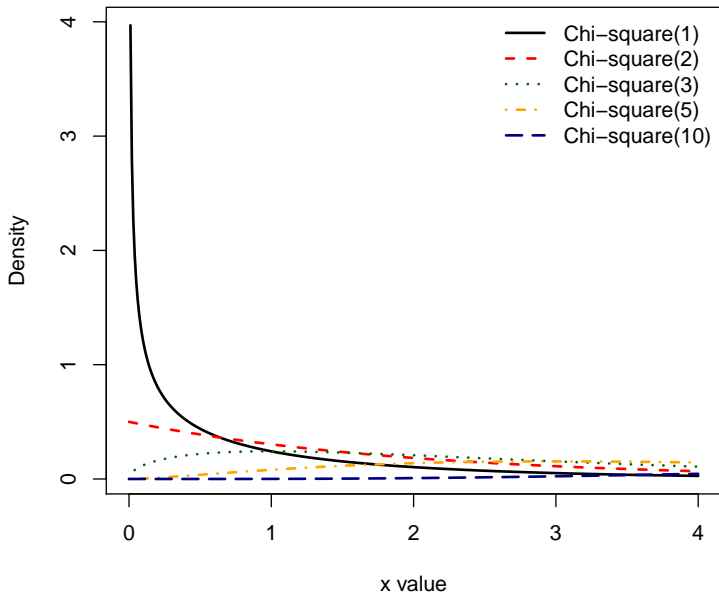
$$\begin{aligned} f(w) &= \frac{1}{2^k \Gamma(k)} w^k \exp\left[-\frac{w}{2}\right] \\ &= \frac{w^{\frac{k-2}{2}} \exp\left(-\frac{w}{2}\right)}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} \end{aligned}$$

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

For $W \sim \chi_k^2$:

$$E(W) = k$$

and:

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

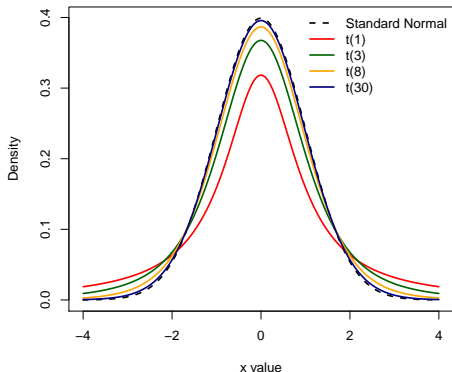
Also:

- 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:

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



Student's t Characteristics

- Mean: $\mu = 0$
- Variance: $\sigma^2 = \frac{k}{k-2}$
- Skewness is 0 for $k > 3$; otherwise undefined
- $t_k \rightarrow N(0, 1)$ as $k \rightarrow \infty$
- If if $Z \sim N(0, 1)$, $W \sim \chi_k^2$, 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:

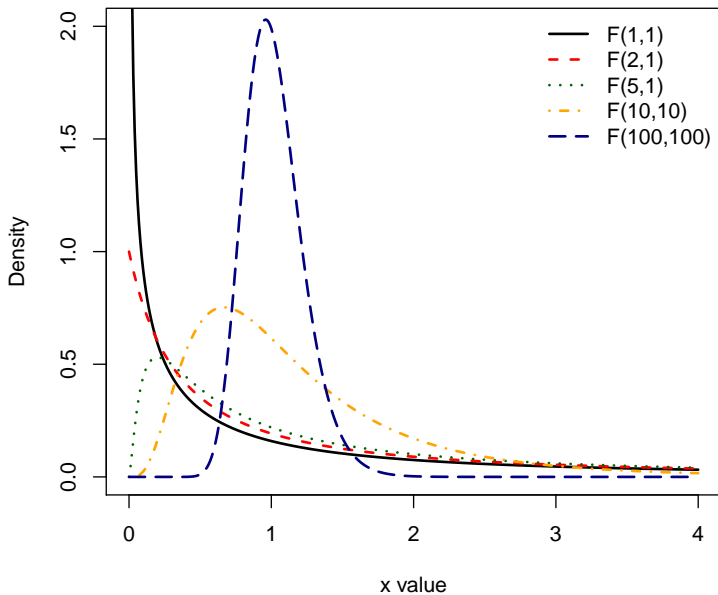
$$f(x) = \frac{\left(\frac{kx}{kx+\ell}\right)^{k/2} \left(1 - \frac{kx}{kx+\ell}\right)^{\ell/2}}{x B(k/2, \ell/2)}$$

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

And we say:

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

F Densities

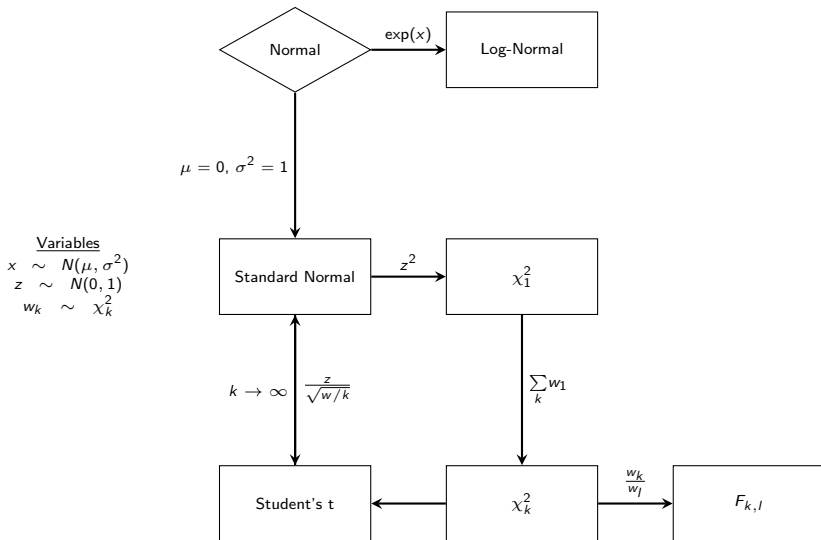


For an $F_{k,\ell}$ distribution:

- $E(X) = \frac{\ell}{\ell-2}$
- $\text{Var}(X) = \frac{2\ell^2(k+\ell-2)}{k(\ell-2)^2(\ell-4)}$
- $\text{Skewness} = \frac{(2k+\ell-2)\sqrt{8(\ell-4)}}{(\ell-6)\sqrt{k(k+\ell-2)}}$
- For independent $W_1 \sim \chi_k^2$ and $W_2 \sim \chi_\ell^2$:

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

Relationships Among Continuous Distributions



Summary

A Few Distributions

Distribution	First Parameter	Meaning	Second Parameter	Meaning
Bernoulli	π	Pr(Success)		
Binomial	n	N of trials	π	Pr(Success)
Geometric	π	Pr(Success)		
Negative Binomial	r	Rank of "success"	π	Pr(Success)
Poisson	λ	"Arrival rate"		
Uniform	a	Minimum	b	Maximum
Normal	μ	Mean	σ^2	Variance
Lognormal	m	Mean	s^2	Variance
Student's t	k	Degrees of freedom		
Chi-Square	k	Degrees of freedom		
F	k	Degrees of freedom	ℓ	Degrees of freedom

Other Useful Distributions

There are a lot; here, in no particular order:

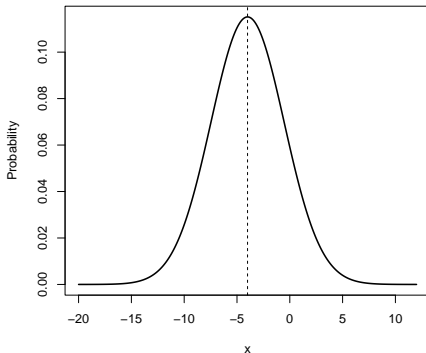
- The **beta distribution** – a continuous two-parameter distribution defined on $[0, 1]$; widely used to model percentages / proportions, and as a “prior” in Bayesian analysis.
- The **exponential distribution** – a continuous one-parameter distribution defined on $[0, \infty)$ that is related to the Poisson distribution (and others).
- The **Weibull distribution** – a continuous two parameter distribution defined on $[0, \infty)$ widely used in survival analysis.
- The **gamma distribution** – a continuous two parameter distribution defined on $[0, \infty)$ that encompasses the exponential, chi-square, and other distributions as special cases.
- The **logistic distribution** – a continuous two parameter distribution defined on \mathbb{R} that resembles the Normal and is widely used in statistics (e.g., logistic regression) and machine learning.
- The **Cauchy distribution** – a continuous two parameter distribution defined on \mathbb{R} and that is interesting mainly because it has no defined MGF.

Practical Things

How To Plot A Density / PDF

"Plot a Normal distribution with a mean of -4 and a variance of 12."

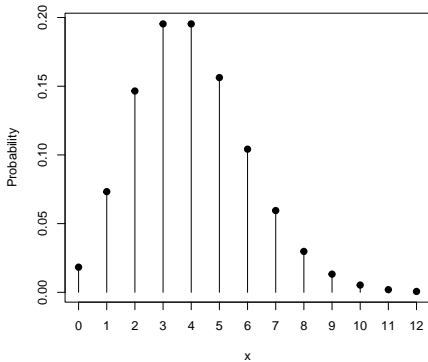
```
> x<-seq(-20,12,by=0.1)  
> PlotNorm<-dnorm(x,-4,sqrt(12))
```



Plotting A Discrete Density

"Plot a Poisson distribution with $\lambda = 4$."

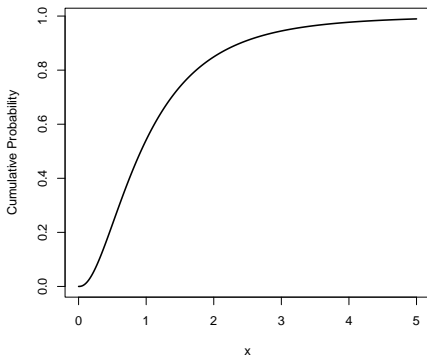
```
> x<-seq(0,12)
> PlotPois<-dpois(x,4)
```



How To Plot A CDF

"Plot the CDF of an F distribution with $k = 5$ and $\ell = 12$."

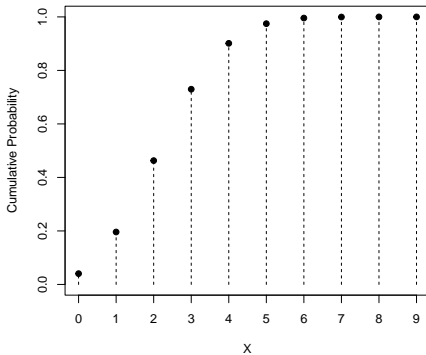
```
> x<-seq(0,5,by=0.01)  
> PlotF<-pf(x,5,12)
```



CDF: Discrete Distribution

"Plot the CDF of a binomial distribution with $\pi = 0.3$ and $n = 9$."

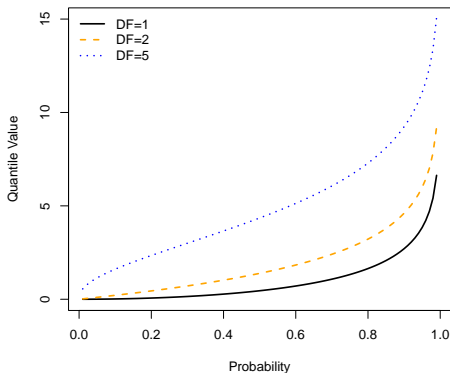
```
> x<-seq(0,9)
> PlotBinom9<-pbinom(x,9,0.3)
```



Quantiles of a Distribution

“Plot the quantiles of three χ^2 distributions with one, two, and five degrees of freedom.”

```
> P<-seq(0.01,0.99,by=0.01) # probabilities  
> ChiSq1<-qchisq(P,1) # df=1  
> ChiSq2<-qchisq(P,2) # df=2  
> ChiSq5<-qchisq(P,5) # df=5
```



Simulating Random Variables

Commands for Generating Random Variates

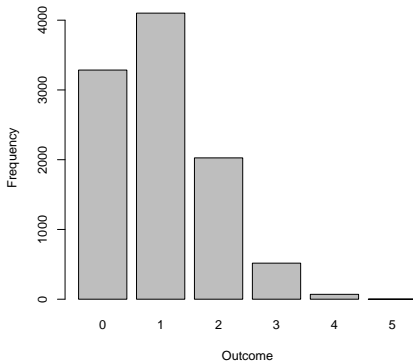
Distribution	R	Stata
Binomial(n, π)	<code>rbinom()</code>	<code>rbinomial</code>
Geometric(π)	<code>rgeom()</code>	(<i>"by hand"</i>)
Negative Binomial(n, π)	<code>rnbinom()</code>	<code>rnbinomial</code>
Poisson(λ)	<code>rpois()</code>	<code>rpoisson</code>
Uniform(0, 1)	<code>runif()</code>	<code>runiform</code>
Normal(0, 1)	<code>rnorm()</code>	<code>rnormal</code>
Lognormal(0, 1)	<code>rlnorm()</code>	<code>xlgn*</code>
Student's $t(k)$	<code>rt()</code>	<code>rt</code>
Chi-Square(k)	<code>rchisq()</code>	<code>rchi2</code>
$F(k, \ell)$	<code>rf()</code>	<code>rndf*</code>

Note: Stata commands marked with an asterisk are from Hilbe's `-rnd-` group of commands.

Drawing From A Distribution

"Draw 10,000 random draws from a binomial distribution with $n = 5$ and $\pi = 0.2$."

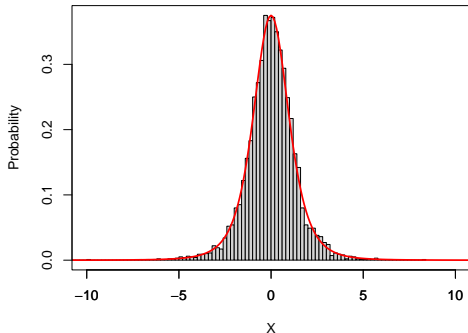
```
> Xbinom5point2<-rbinom(10000,5,0.2)
```



More Draws From A Distribution

“Draw 5000 random draws from a t distribution with 4 degrees of freedom, and compare the distribution of values to the theoretical density.”

```
> TDraws<-rt(5000,4)
```

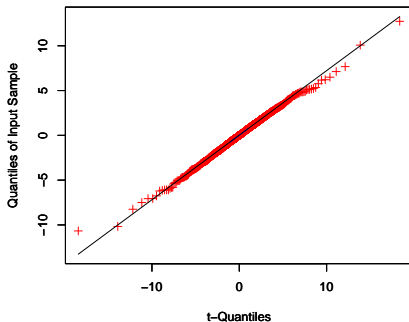


(continued...)

(Same comparison, using a Q-Q plot...)

```
> library(Dowd)
> TQQPlot(TDraws,4)
```

QQ Plot of Sample Data versus Student-t with 4 Degrees of freedom



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
```

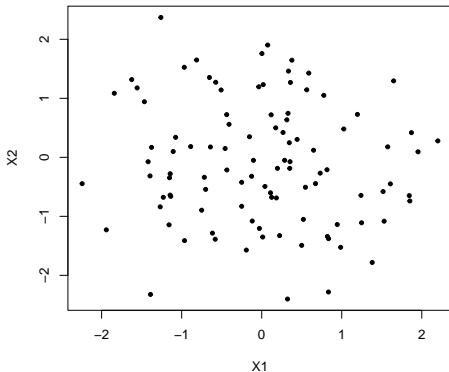
```
> seed<-1077  
> set.seed(seed) # resetting the seed  
> rt(3,1) # different values for the draws  
[1] -0.5211  7.9161 -155.3186
```

```
> seed<-3229 # original seed  
> set.seed(seed)  
> rt(3,1) # identical values of the draws  
[1] -0.1113 -0.7306  1.9839
```

Seeds and Simulations

The right way:

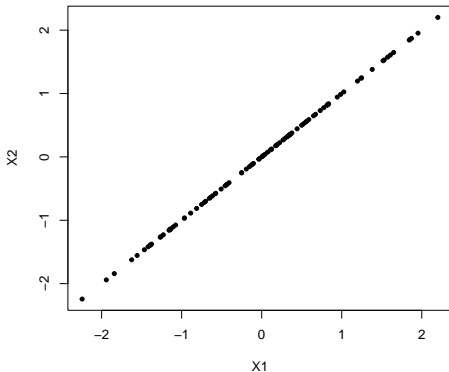
```
> X<-matrix(NA,nrow=100,ncol=2)
> set.seed(7222009)
> for(i in 1:2){
+   X[,i]<-rnorm(100)}
```



Seeds and Simulations (continued)

The wrong way:

```
> X<-matrix(NA,nrow=100,ncol=2)
> for(i in 1:2){
+   set.seed(7222009)
+   X[,i]<-rnorm(100)}
```



In general:

1. **Always** set a seed value.
2. Try to use a single/consistent seed value over time/projects.
3. Keep seeds *outside* loops/apply statements.
6. Use `ddists` and `pdists` for theoretical quantities, `rdists` for generating simulated “data” / variates.
5. Plot discrete distributions discretely....
6. ...and continuous distributions continuously.