Convergence in Distribution
Central Limit Theorem

Statistics 110

Summer 2006

Convergence in Distribution

Theorem. *Let $X \sim Bin(n,p)$ and let $\lambda = np$, Then*

$$\lim_{n\to\infty} P[X=x] = \lim_{n\to\infty} \binom{n}{x} p^x (1-p)^{n-x} = \frac{e^{-\lambda}\lambda^x}{x!}$$

So when n gets large, we can approximate binomial probabilities with Poisson probabilities.

Proof.

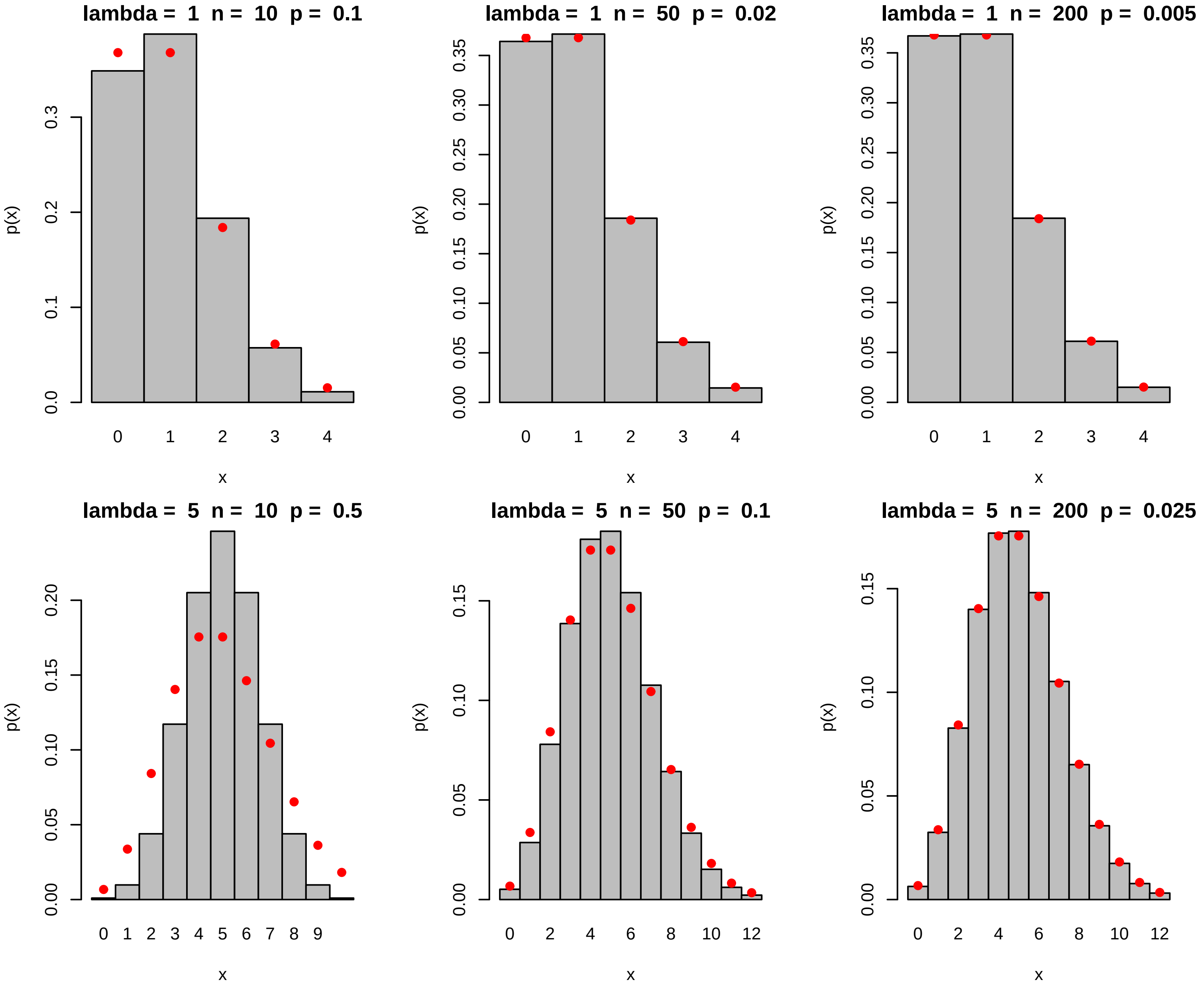
$$\begin{aligned}\lim_{n\to\infty} \binom{n}{x} p^x (1-p)^{n-x} &= \lim_{n\to\infty} \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1-\frac{\lambda}{n}\right)^{n-x} \\ &= \frac{n!}{x!(n-x)!}\lambda^x \left(\frac{1}{n^x}\right) \left(1-\frac{\lambda}{n}\right)^{-x} \left(1-\frac{\lambda}{n}\right)^{n}\end{aligned}$$

$$
\begin{aligned}
&= \frac{n!}{x!(n-x)!}\lambda^x \left(\frac{1}{n^x}\right)\left(1-\frac{\lambda}{n}\right)^{-x}\left(1-\frac{\lambda}{n}\right)^{n} \\
&= \frac{\lambda^x}{x!}\lim_{n\to\infty} \underbrace{\frac{n!}{(n-x)!}\frac{1}{(n-\lambda)^x}}_{\to 1} \quad \underbrace{\left(1-\frac{\lambda}{n}\right)^{n}}_{\to e^{-\lambda}} \\
&= \frac{e^{-\lambda}\lambda^x}{x!}
\end{aligned}
$$

□

Note that approximation works better when n is large and p is small as can been seen in the following plot. If p is relatively large, a different approximation should be used. This is coming later.

(Note in the plot, bars correspond to the true binomial probabilities and the red circles correspond to the Poisson approximation.)


lambda = 1 n = 10 p = 0.1
lambda = 1 n = 50 p = 0.02
lambda = 1 n = 200 p = 0.005
lambda = 5 n = 10 p = 0.5
lambda = 5 n = 50 p = 0.1
lambda = 5 n = 200 p = 0.025
p(x)
x
0.0
0.1
0.2
0.3
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0
1
2
3
4
5
6
7
8
9
10
12

Example: Let $Y_1, Y_2, \ldots$ be iid $Exp(1)$. Then

$$X_n = Y_1 + Y_2 + \ldots + Y_n \sim Gamma(n, 1)$$

which has

$$E[X_n] = n; \qquad \mathrm{Var}(X_n) = n; \qquad SD(X_n) = \sqrt{n}$$

Thus $Z_n = \frac{X_n - n}{\sqrt{n}}$ has mean $= 0$ and variance $= 1$.

Lets compare its distribution to $Z \sim N(0, 1)$. i.e. Is

$$P[-1 \leq Z_n \leq 2] \approx P[-1 \leq Z \leq 2]?$$

Let

$$Z_n = \frac{X_n - n}{\sqrt{n}}; \qquad X_n = n + \sqrt{n} Z_n$$

$$f_{Z_n}(z) = f_{X_n}(n + \sqrt{n} z) \times \sqrt{n}$$

$$
\begin{aligned}
P[a \le Z_n \le b] &= \int_a^b f_{Z_n}(z)dz \\
&= \int_a^b \sqrt{n} f_{X_n}(n + \sqrt{n}z)dz \\
&= \int_a^b \sqrt{n} \frac{(n + \sqrt{n}z)^{n-1}}{(n-1)!} e^{-(n+\sqrt{n}z)} dz
\end{aligned}
$$

To go further we need Stirling's Formula: $n! \approx n^n e^{-n} \sqrt{2\pi n}$. So

$$
\begin{aligned}
f_{X_n}(n + \sqrt{n}z)\sqrt{n} &= e^{-n-z\sqrt{n}}(n + z\sqrt{n})^{n-1} \frac{\sqrt{n}}{(n-1)!} \\
&\approx \frac{e^{-n-z\sqrt{n}}(n + z\sqrt{n})^{n-1}\sqrt{n}}{(n-1)^{n-1}e^{-n+1}\sqrt{2\pi n}} \\
&\approx \frac{1}{\sqrt{2\pi}} \underbrace{e^{-z\sqrt{n}} \left(1 + \frac{z}{\sqrt{n}}\right)^n}_{g_n(z)}
\end{aligned}
$$

$$
\begin{aligned}
\log(g_n(z)) &= -z\sqrt{n} + n\log\left(1+\frac{z}{\sqrt{n}}\right) \\
&= -z\sqrt{n} + n\left[\frac{z}{\sqrt{n}} - \frac{1}{2}\frac{z^2}{n} + \frac{1}{3}\frac{z^3}{n^{3/2}} - \dots\right] \approx -\frac{1}{2}z^2 + O\left(\frac{1}{\sqrt{n}}\right)
\end{aligned}
$$

so

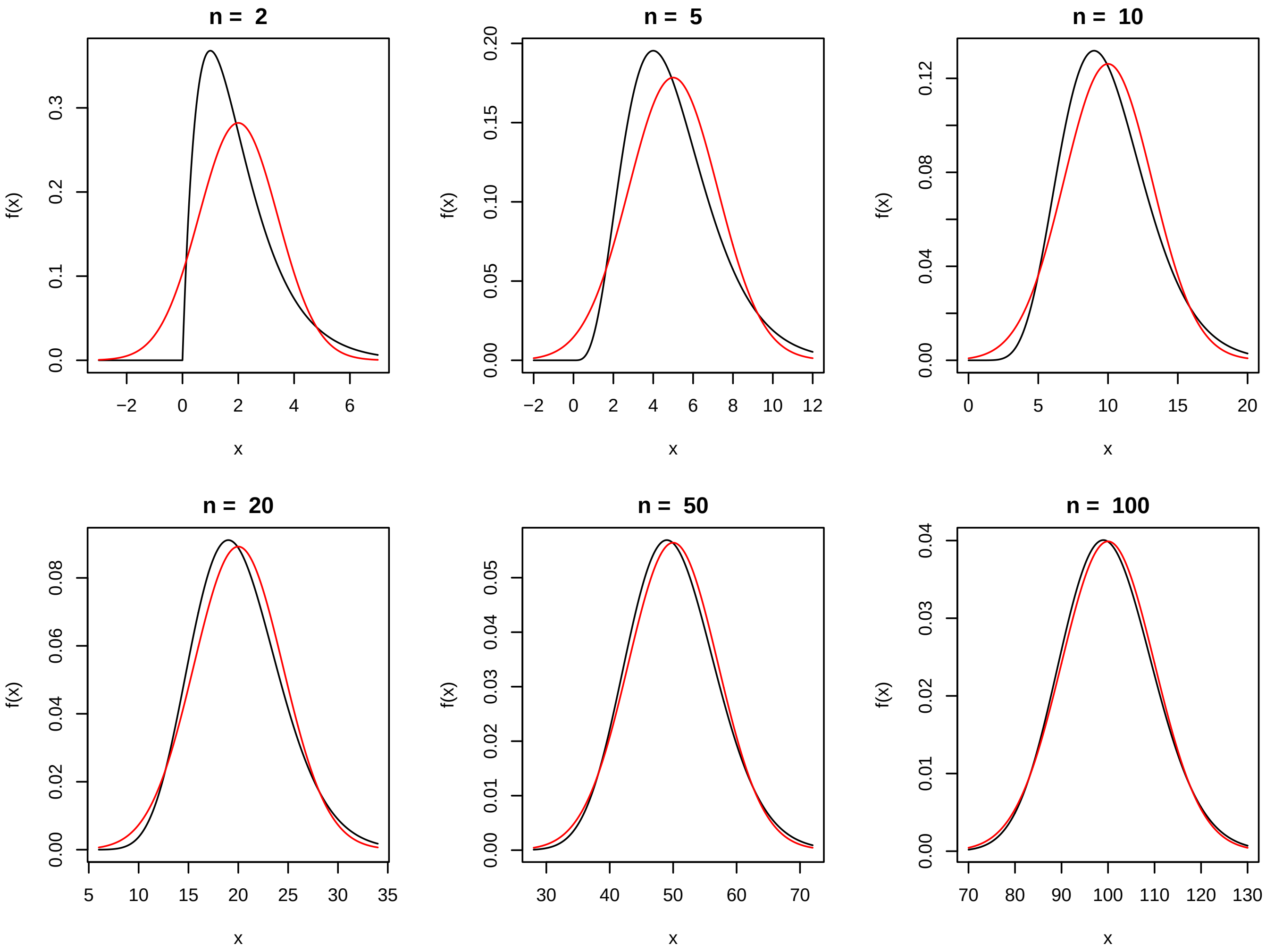
$$f_{X_n}(n + z\sqrt{n})\sqrt{n} \approx \frac{1}{\sqrt{2\pi}}e^{-z^2/2}$$

Thus

$$P[a \leq Z_n \leq b] \to \int_a^b \frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz = P[a \leq Z \leq b]$$

So as n increases, the distribution of Z_n gets closer and closer to a $N(0,1)$.

Another way of thinking of this, is that the distribution of $X_n = n + Z_n\sqrt{n}$ approaches that of a $N(n, n)$.



Definition. *Let $X_1, X_2, \ldots$ be a sequence of RVs with cumulative distribution functions $F_1, F_2, \ldots$ and let X be a RV with distribution F. We say X_n* **Converges in Distribution** *to X if*

$$\lim_{n \to \infty} F_n(x) = F(x)$$

at every point at which F is continuous. $X_n \xrightarrow{\mathcal{D}} X$

An equivalent statement to this is that for all a and b where F is continuous

$$P[a \leq X_n \leq b] \to P[a \leq X \leq b]$$

Note that if X_n and X are discrete distributions, this condition reduces to $P[X_n = x_i] \to P[X = x_i]$ for all support points x_i.

Note that an equivalent definition of convergence in distribution is that $X_n \xrightarrow{\mathcal{D}} X$ if $E[g(X_n)] \to E[g(X)]$ for all bounded, continuous functions $g(\cdot)$.

This statement of convergence in distribution is needed to help prove the following theorem

Theorem. [Continuity Theorem] *Let X_n be a sequence of random variables with cumulative distribution functions $F_n(x)$ and corresponding moment generating functions $M_n(t)$. Let X be a random variable with cumulative distribution function $F(x)$ and moment generating function $M(t)$. If $M_n(t) \to M(t)$ for all t in an open interval containing zero, then $F_n(x) \to F(x)$ at all continuity points of F. That is $X_n \xrightarrow{\mathcal{D}} X$.*

Thus the previous two examples (Binomial/Poisson and Gamma/Normal) could be proved this way.

For the Gamma/Normal example

$$M_{Z_n}(t) = M_{X_n}\left(\frac{t}{\sqrt{n}}\right) e^{-t\sqrt{n}} = \left(\frac{1}{1-\frac{t}{\sqrt{n}}}\right)^n e^{-t\sqrt{n}}$$

Similarly to the earlier proof, its easier to work with $\log M_{Z_n}(t)$

$$\begin{aligned}\log M_{Z_n}(t) &= -t\sqrt{n} - n\log\left(1-\frac{t}{\sqrt{n}}\right) \\ &= -t\sqrt{n} - n\left[-\frac{t}{\sqrt{n}} - \frac{1}{2}\frac{t^2}{n} - \frac{1}{3}\frac{t^3}{n^{3/2}} - \ldots\right] \\ &= \frac{1}{2}t^2 + O\left(\frac{1}{\sqrt{n}}\right)\end{aligned}$$

Thus

$$M_{Z_n}(t) \to e^{t^2/2}$$

which is the MGF for a standard normal.

Central Limit Theorem

Theorem. [Central Limit Theorem (CLT)] *Let $X_1, X_2, X_3, \ldots$ be a sequence of independent RVs having mean μ and variance σ^2 and a common distribution function $F(x)$ and moment generating function $M(t)$ defined in a neighbourhood of zero. Let*

$$S_n = \sum_{i=1}^{n} X_n$$

Then

$$\lim_{n\to\infty} P\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right] = \Phi(x)$$

That is

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{\mathcal{D}} N(0,1)$$

Proof. Without a loss of generality, we can assume that $\mu = 0$. So let $Z_n = \frac{S_n}{\sigma\sqrt{n}}$. Since S_n is the sum of n iid RVs,

$$M_{S_n}(t) = (M(t))^n; \qquad M_{Z_n}(t) = \left(M\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n$$

Taking a Taylor series expansion of $M(t)$ around 0 gives

$$M(t) = M(0) + M'(0)t + \frac{1}{2}M''(0)t^2 + \epsilon_t = 1 + \frac{1}{2}\sigma^2 t^2 + O(t^3)$$

since $M(0) = 1, M'(0) = \mu = 0, M''(0) = \sigma^2$. So

$$\begin{aligned} M\left(\frac{t}{\sigma\sqrt{n}}\right) &= 1 + \frac{1}{2}\sigma^2\left(\frac{t}{\sigma\sqrt{n}}\right)^2 + O\left(\left(\frac{t}{\sigma\sqrt{n}}\right)^3\right) \\ &= 1 + \frac{t^2}{2n} + O\left(\frac{1}{n^{3/2}}\right) \end{aligned}$$

This gives

$$M_{Z_n}(t) = \left(1 + \frac{t^2}{2n} + O\left(\frac{1}{n^{3/2}}\right)\right)^n \to e^{t^2/2}$$

□

Note that the requirement of a MGF is not needed for the theorem to hold. In fact, all that is needed is that $\mathrm{Var}(X_i) = \sigma^2 < \infty$. A standard proof of this more general theorem uses the characteristic function (which is defined for any distribution)

$$\phi(t) = \int_{-\infty}^{\infty} e^{itx} f(x)dx = M(it)$$

instead of the moment generating function $M(t)$, where $i = \sqrt{-1}$.

Thus the CLT holds for distributions such as the log normal, even though it doesn't have a MGF.

Also, the CLT is often presented in the following equivalent form

$$Z_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \sqrt{n}\frac{\bar{X}_n - \mu}{\sigma} \xrightarrow{\mathcal{D}} N(0,1)$$

To see this is the same, just multiply the numerator and denominator by n in the first form to get the statement about S_n.

The common way that this is used is that

$$S_n \overset{approx.}{\sim} N\left(n\mu, n\sigma^2\right) \quad \text{or} \quad \bar{X}_n \overset{approx.}{\sim} N\left(\mu, \frac{\sigma^2}{n}\right)$$

Example: Insurance claims

Suppose that an insurance company has 10,000 policy holders. The expected yearly claim per policyholder is $240 with a standard deviation of $800. What is the approximate probability that the total yearly claims $S_{10,000} >$ $2.6 Million

$$E[S_{10,000}] = 10,000 \times 240 = 2,400,000$$
$$\mathrm{SD}(S_{10,000}) = \sqrt{10,000} \times 800 = 80,000$$

$$\begin{aligned} P[S_{10,000} > 2,600,000] \\ &= P\left[\frac{S_{10,000} - 2,400,000}{80,000} > \frac{2,600,000 - 2,400,000}{80,000}\right] \\ &\approx P[Z > 2.5] = 0.0062 \end{aligned}$$

Note that this probability statement does not use anything about the distribution of the original policy claims except their mean and standard deviation. Its probable that their distribution is highly skewed right (since $\mu_x << \sigma_x$), but the calculations ignore this fact.

One consequence of the CLT is the normal approximation to the binomial. If $X_n \sim Bin(n,p)$ and $\hat{p}_n = \frac{X_n}{n}$, then (since X_n can be thought of the sum of n Bernoulli's)

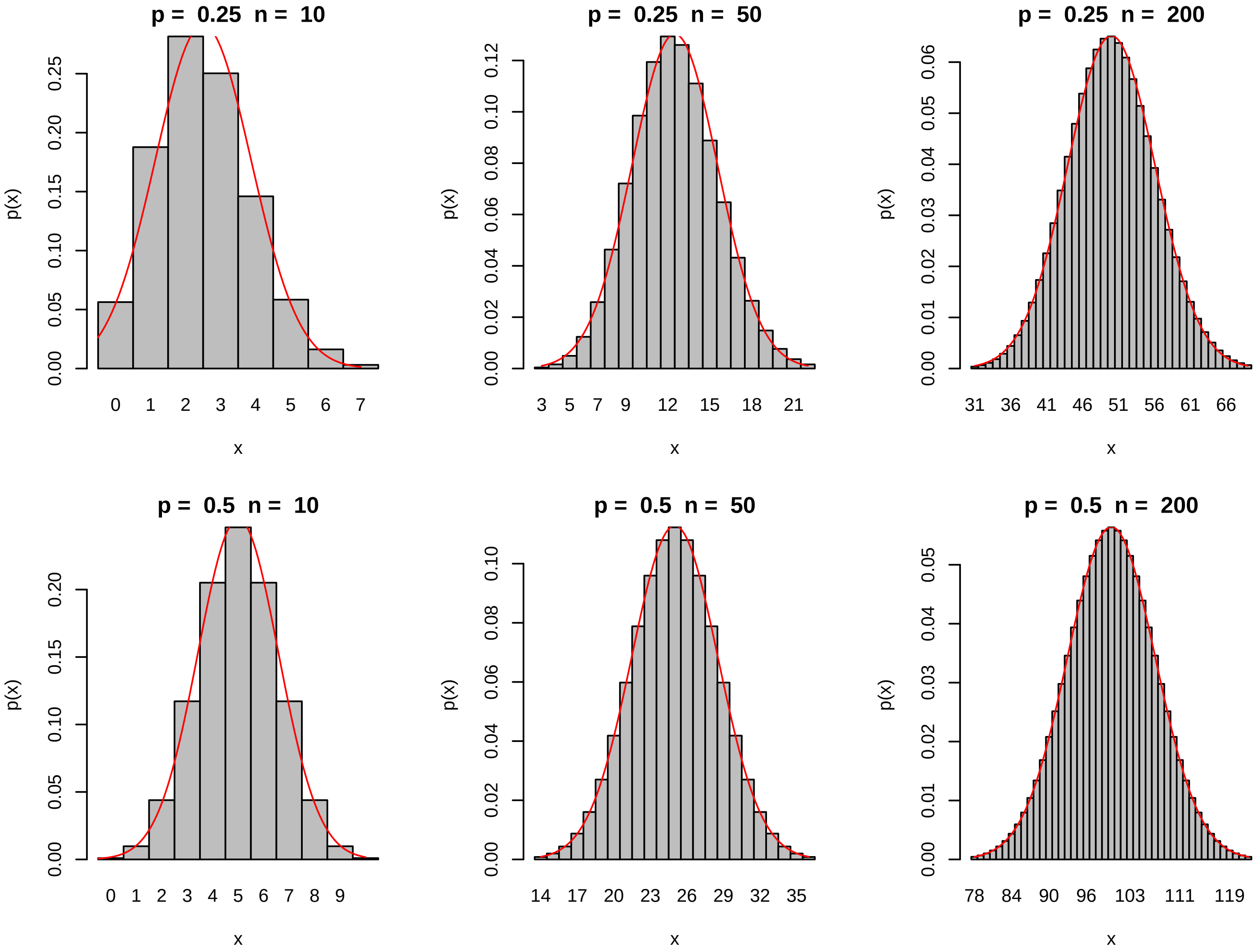
$$\frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow{\mathcal{D}} N(0,1); \qquad \frac{\hat{p}_n - p}{\sqrt{p(1-p)/n}} \xrightarrow{\mathcal{D}} N(0,1)$$

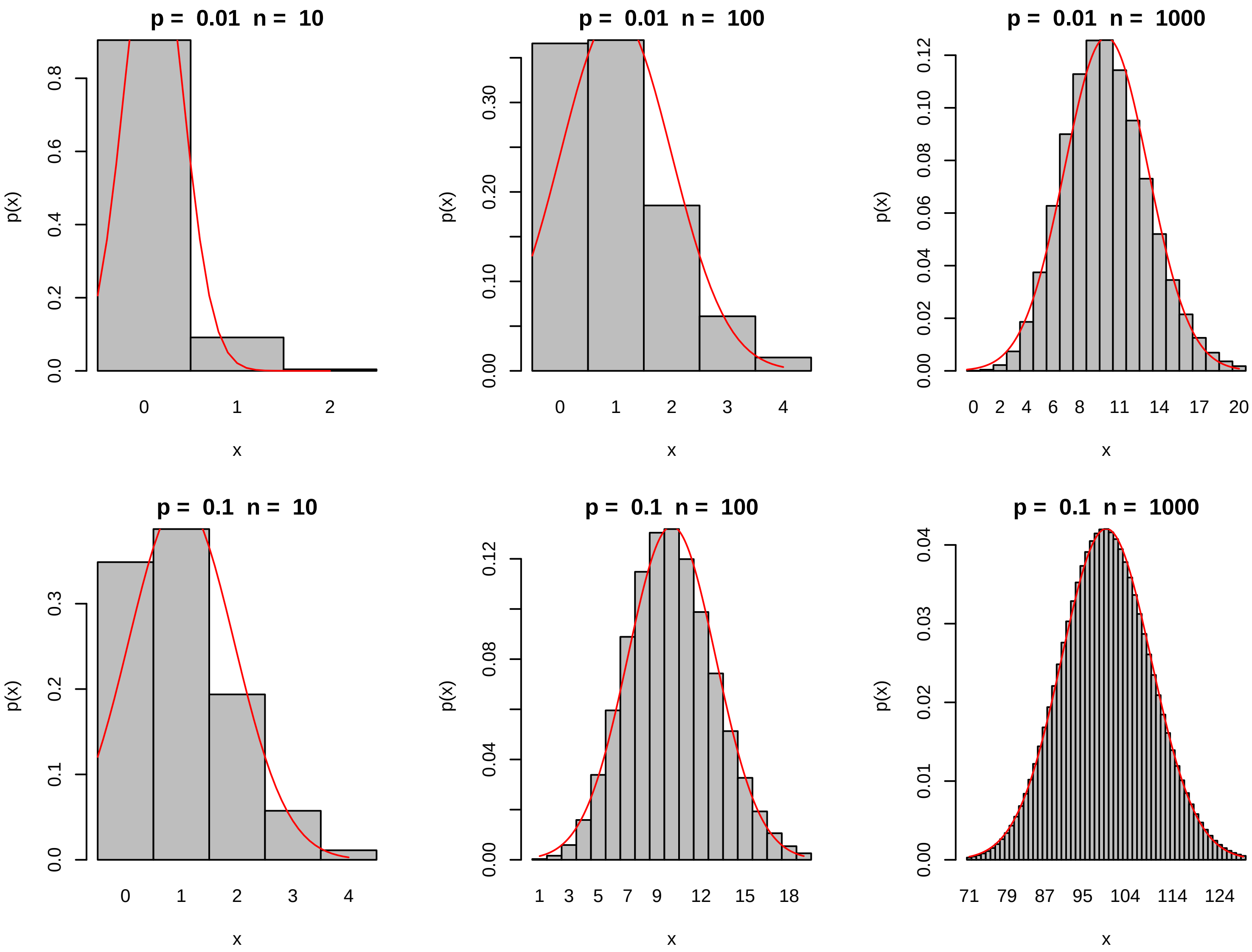
Another way of think of this is that

$$X_n \overset{approx.}{\sim} N(np, np(1-p)); \qquad \hat{p}_n \overset{approx.}{\sim} N\left(p, \frac{p(1-p)}{n}\right)$$

This approximation works better when p is closer to $\frac{1}{2}$ than when p is near 0 or 1.

A rule of thumb is that is ok to use the normal approximation when $np \geq 5$ and $n(1-p) \geq 5$ (expect at least 5 successes and 5 failures). (Other books sometimes suggest other values, with the most popular alternative being 10.)


p = 0.25 n = 10
p = 0.25 n = 50
p = 0.25 n = 200
p = 0.5 n = 10
p = 0.5 n = 50
p = 0.5 n = 200
p(x)
x
0.00
0.05
0.10
0.15
0.20
0.25
0 1 2 3 4 5 6 7
0.00
0.02
0.04
0.06
0.08
0.10
0.12
3 5 7 9 12 15 18 21
0.00
0.01
0.02
0.03
0.04
0.05
0.06
31 36 41 46 51 56 61 66
0 1 2 3 4 5 6 7 8 9
14 17 20 23 26 29 32 35
78 84 90 96 103 111 119


p = 0.01 n = 10
p(x)
0.0
0.2
0.4
0.6
0.8
0
1
2
x
p = 0.01 n = 100
p(x)
0.00
0.10
0.20
0.30
0
1
2
3
4
x
p = 0.01 n = 1000
p(x)
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0
2
4
6
8
11
14
17
20
x
p = 0.1 n = 10
p(x)
0.0
0.1
0.2
0.3
0
1
2
3
4
x
p = 0.1 n = 100
p(x)
0.00
0.04
0.08
0.12
1
3
5
7
9
12
15
18
x
p = 0.1 n = 1000
p(x)
0.00
0.01
0.02
0.03
0.04
71
79
87
95
104
114
124
x

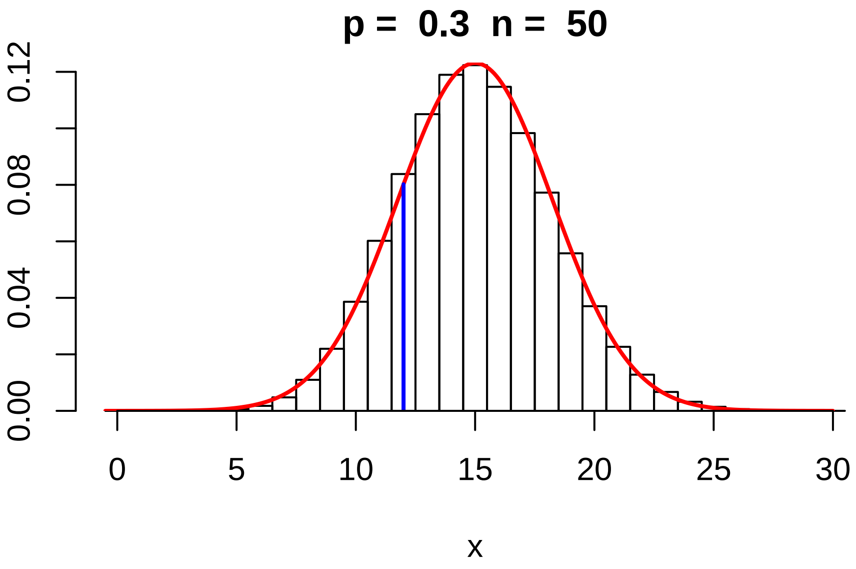
Continuity correction to the binomial approximation

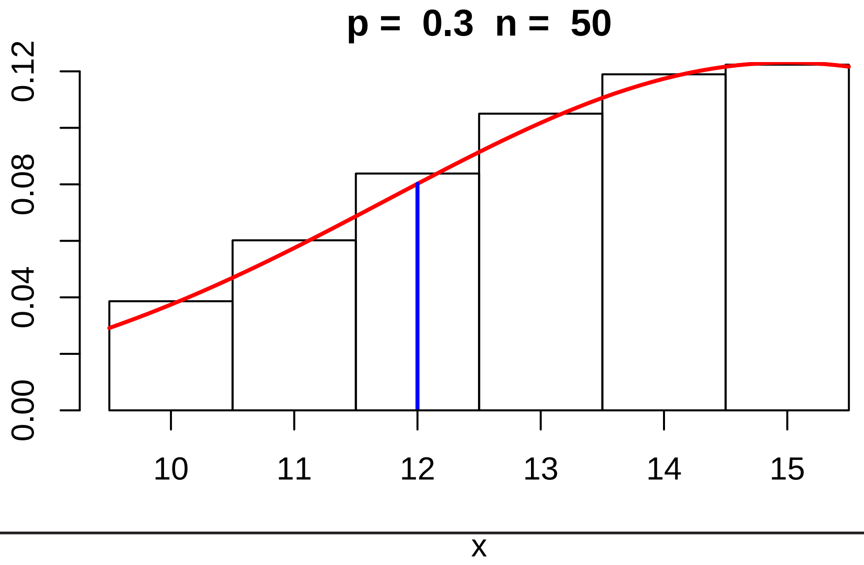
Suppose that $X \sim Bin(50, 0.3)$ and we are interested in

$$P[\hat{p} \leq 0.24] = P[X \leq 12]$$

Notice that the bar corresponding to $X = 12$, the normal curve only picks up about half the area, as the bar actually goes from 11.5 to 12.5.

The normal approximation can be improved if we ask for the area under the normal curve up to 12.5.



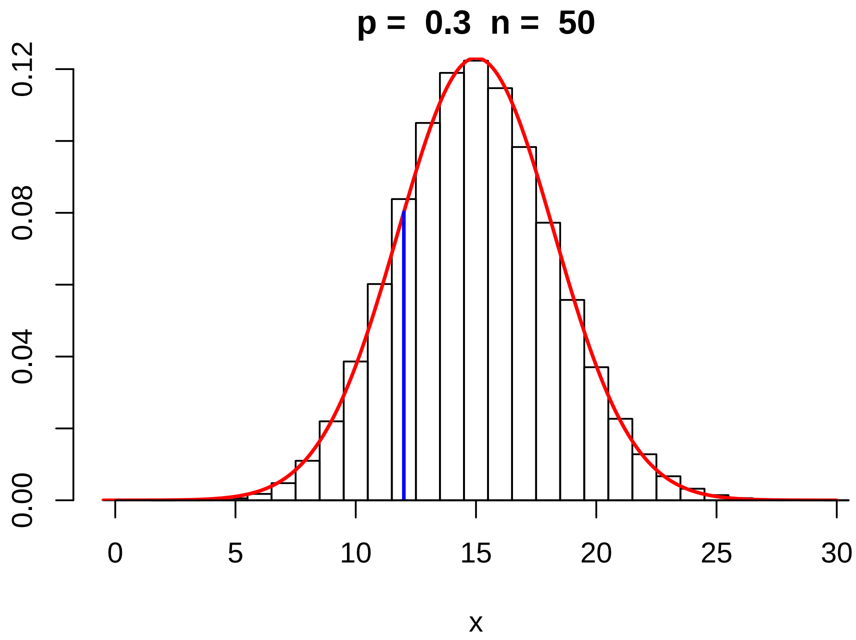


Let $Y \sim N(15, 10.5)$ (approximating normal). Then

$$P[X \leq 12] = 0.2229 \qquad \text{(True Probability)}$$

$$P[Y \leq 12] = 0.1773 \qquad \text{(No correction)}$$

$$P[Y \leq 12.5] = 0.2202 \qquad \text{(With correction)}$$

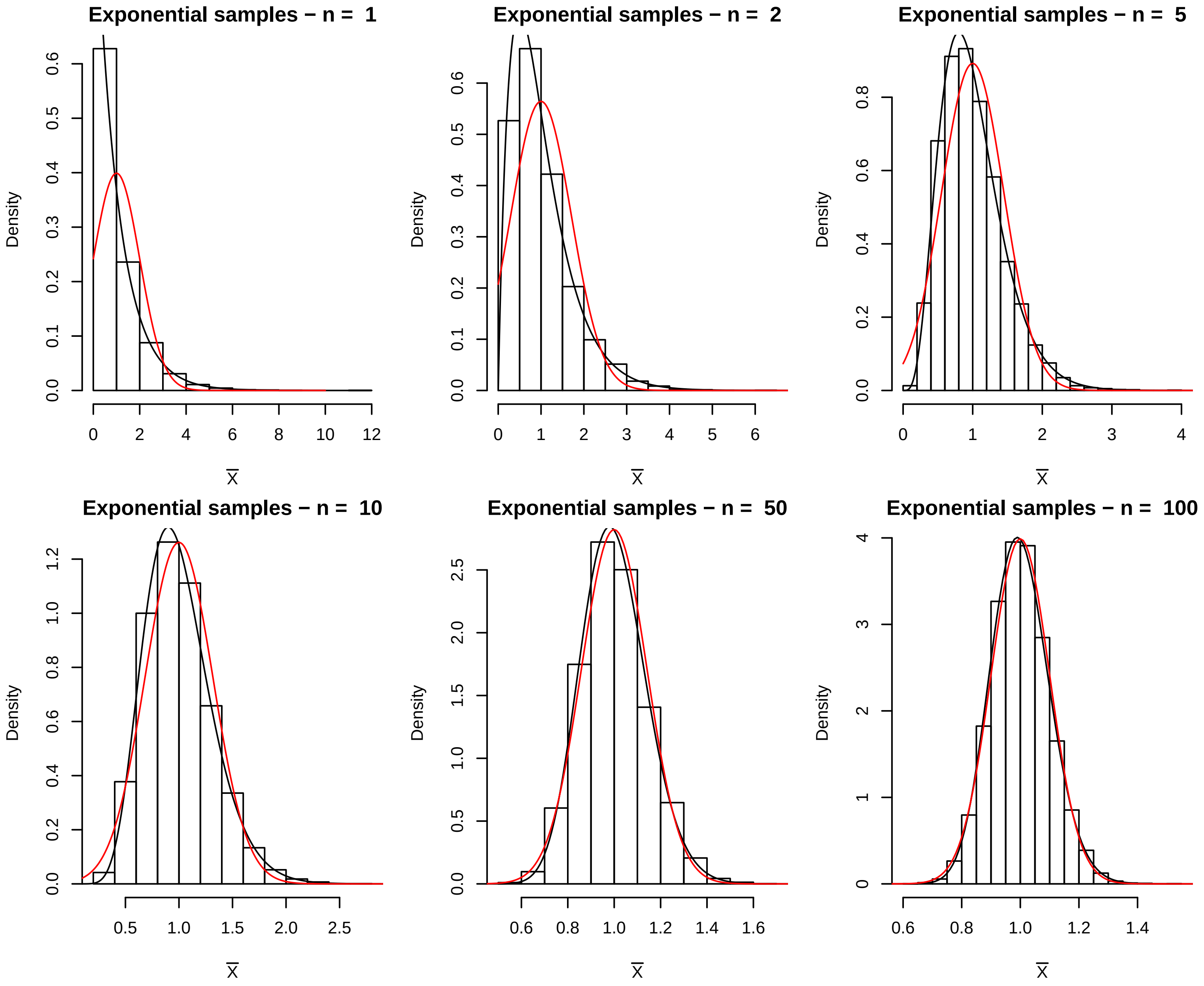


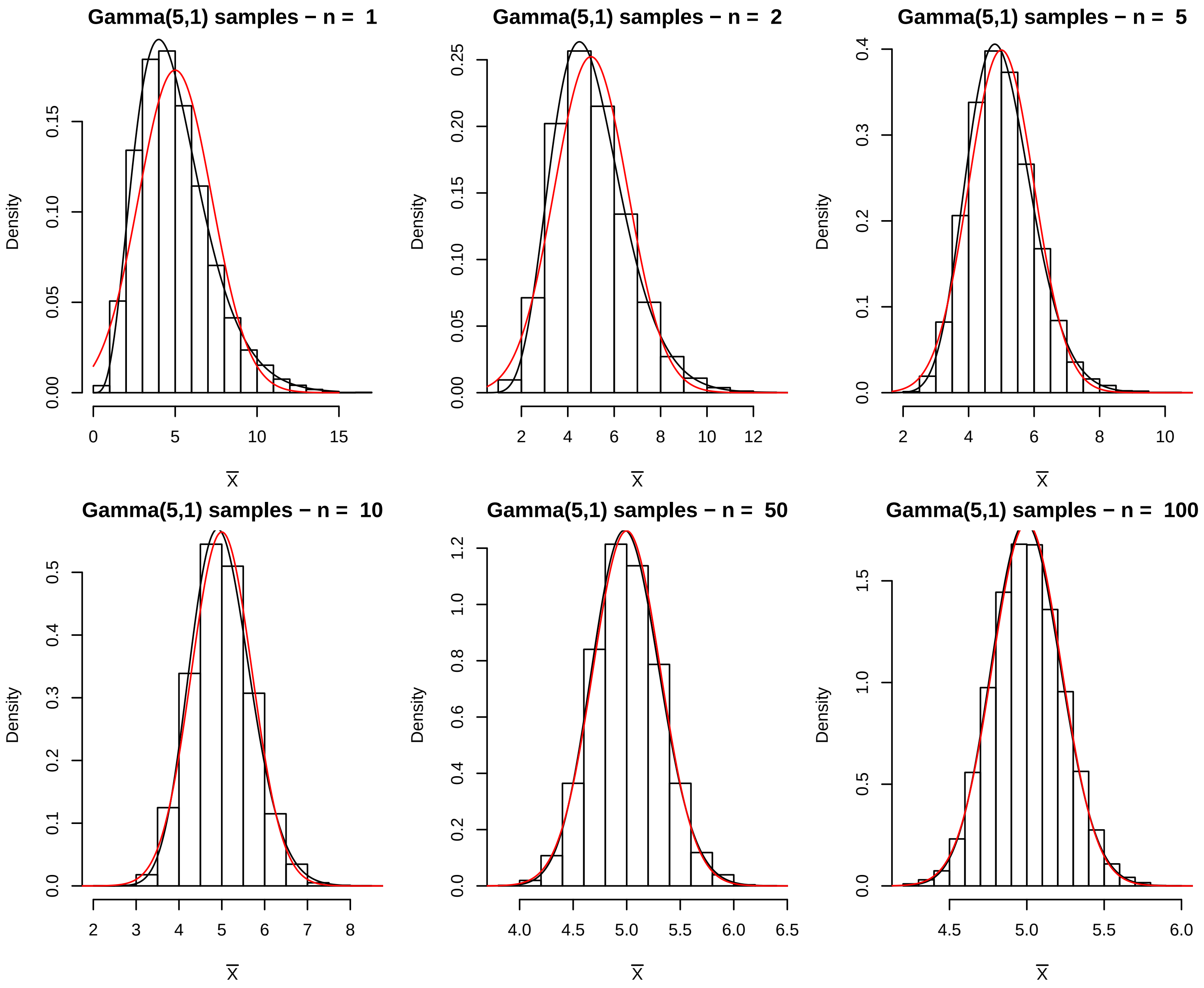
While this does give a better answer for many problems, normally I recommend ignoring it. If the correction makes a difference, you probably want to be doing an exact probability calculation instead.

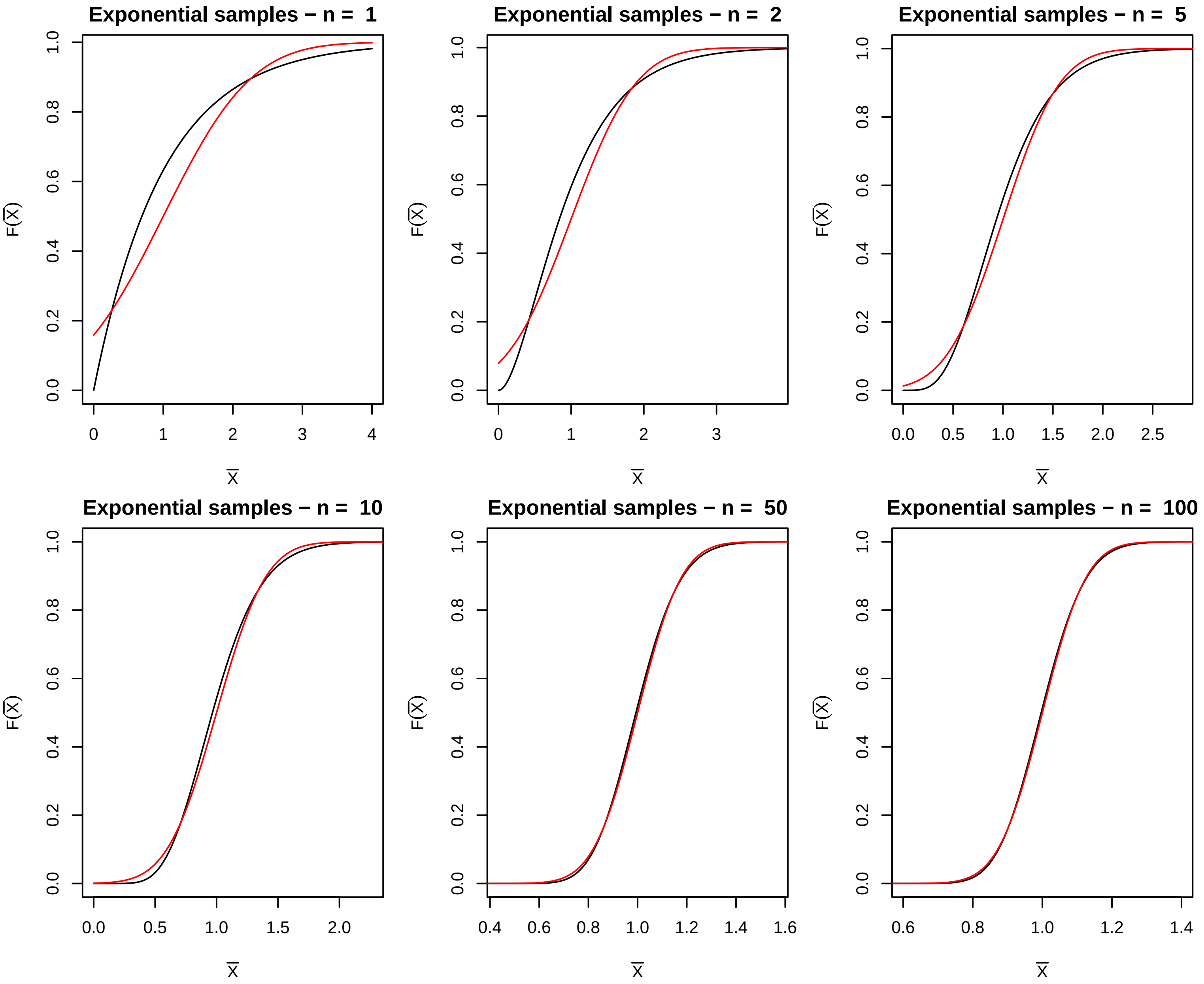
When will the CLT work better?

- Big n
- Distribution of X_i close to normal. Approximation holds exactly if $n = 1$ if $X_i \sim N(\mu, \sigma^2)$.
- X_i roughly symmetric. As we observed with the binomial examples, the closer p was to 0.5, thus closer to symmetry, the better the approximation works. The more skewness there is in the distribution of the observations, the bigger n needs to be.

In the following plots, the histogram represents 10,000 simulated $\bar{X}$s, the black curves are the true densities or CDFs, and the red curves are the normal approximations.


Exponential samples − n = 1
Density
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
2
4
6
8
10
12
$\overline{X}$
Exponential samples − n = 2
Density
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
1
2
3
4
5
6
$\overline{X}$
Exponential samples − n = 5
Density
0.0
0.2
0.4
0.6
0.8
0
1
2
3
4
$\overline{X}$
Exponential samples − n = 10
Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
0.5
1.0
1.5
2.0
2.5
$\overline{X}$
Exponential samples − n = 50
Density
0.0
0.5
1.0
1.5
2.0
2.5
0.6
0.8
1.0
1.2
1.4
1.6
$\overline{X}$
Exponential samples − n = 100
Density
0
1
2
3
4
0.6
0.8
1.0
1.2
1.4
$\overline{X}$


Gamma(5,1) samples − n = 1
Density
0.00
0.05
0.10
0.15
0
5
10
15
X̄
Gamma(5,1) samples − n = 2
Density
0.00
0.05
0.10
0.15
0.20
0.25
2
4
6
8
10
12
X̄
Gamma(5,1) samples − n = 5
Density
0.0
0.1
0.2
0.3
0.4
2
4
6
8
10
X̄
Gamma(5,1) samples − n = 10
Density
0.0
0.1
0.2
0.3
0.4
0.5
2
3
4
5
6
7
8
X̄
Gamma(5,1) samples − n = 50
Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
4.0
4.5
5.0
5.5
6.0
6.5
X̄
Gamma(5,1) samples − n = 100
Density
0.0
0.5
1.0
1.5
4.5
5.0
5.5
6.0
X̄



Exponential samples − n = 1

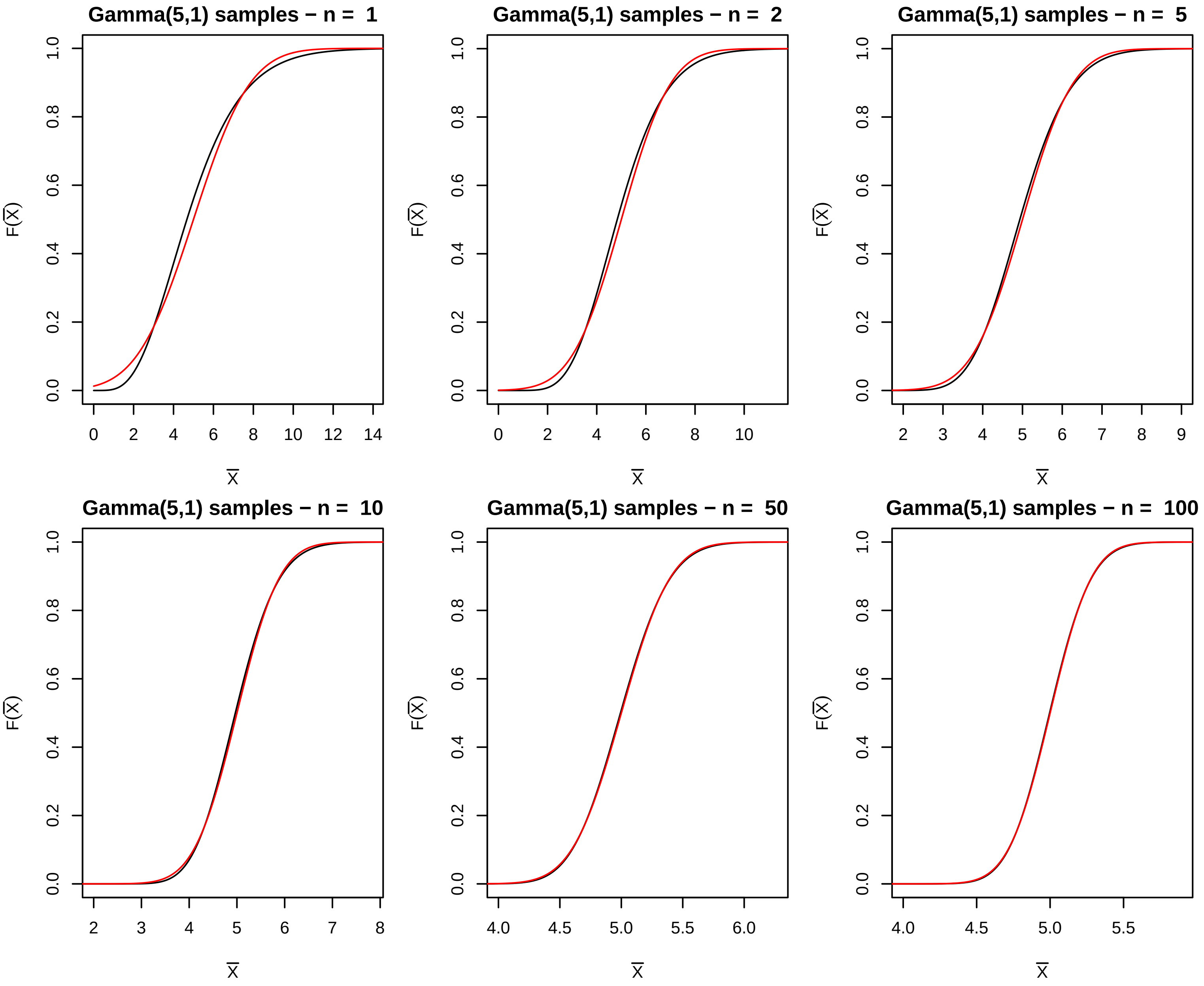
Exponential samples − n = 2

Exponential samples − n = 5

Exponential samples − n = 10

Exponential samples − n = 50

Exponential samples − n = 100


Gamma(5,1) samples − n = 1
Gamma(5,1) samples − n = 2
Gamma(5,1) samples − n = 5
Gamma(5,1) samples − n = 10
Gamma(5,1) samples − n = 50
Gamma(5,1) samples − n = 100
F($\overline{X}$)
$\overline{X}$
0.0
0.2
0.4
0.6
0.8
1.0
0
2
4
6
8
10
12
14
2
3
4
5
6
7
8
9
4.0
4.5
5.0
5.5
6.0

There are other forms of the CLT, which relax the assumptions about the distribution. One example is,

Theorem. [Liapunov's CLT] *Let* $X_1, X_2, \ldots$ *be independent random variables with* $E[X_i] = \mu_i$, $\mathrm{Var}(X_i) = \sigma_i^2$, *and* $E[|X_i - \mu|] = \beta_i$. *Let*

$$B_n = \left(\sum_{i=1}^{n} \beta_i \right)^{1/3} \qquad c_n = \left(\sum_{i=1}^{n} \sigma_i^2 \right)^{1/2} = \mathrm{SD}\left(\sum_{i=1}^{n} X_i \right).$$

Then

$$Y_n = \frac{\sum_{i=1}^{n} (X_i - \mu_i)}{c_n} \xrightarrow{\mathcal{D}} Z \sim N(0,1)$$

if $\frac{B_n}{c_n} \rightarrow 0$

Proof. Omitted □

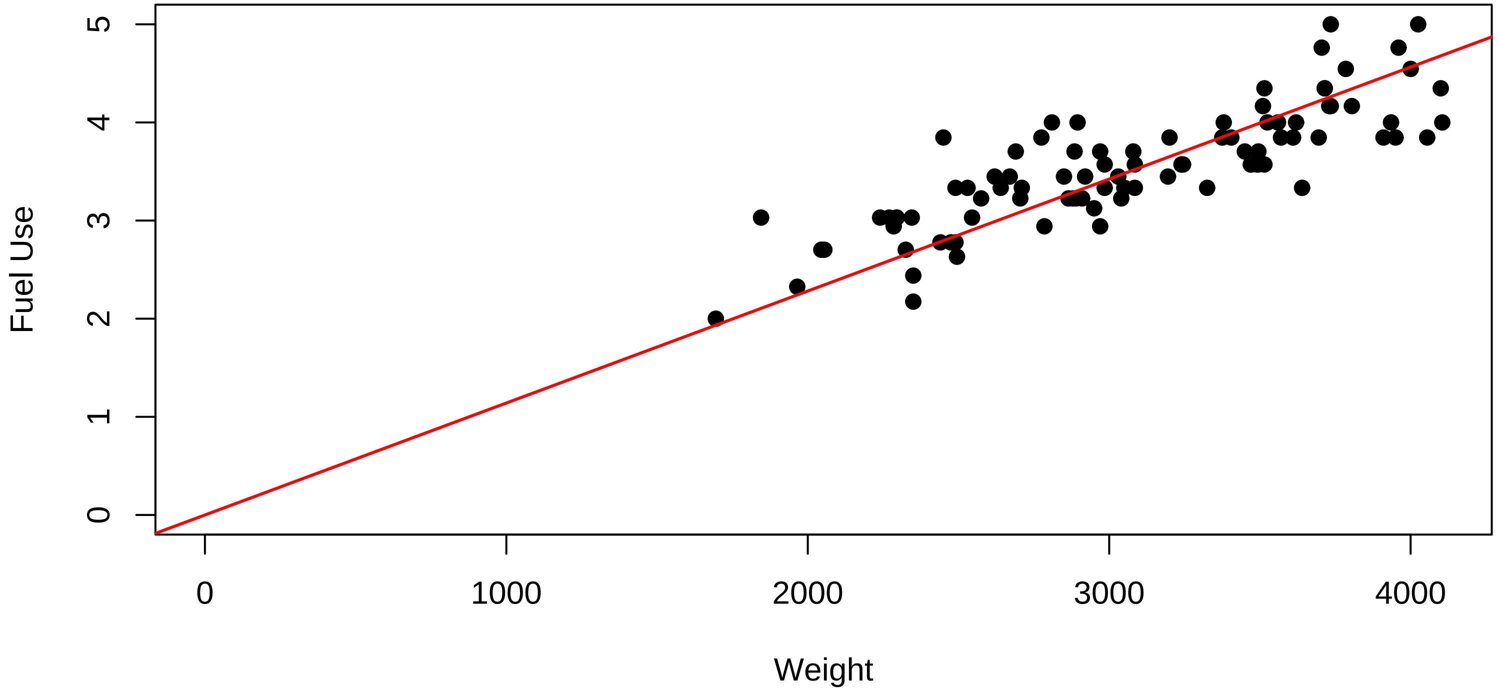
The condition involving B_i and c_i has to do with each term in the sum having roughly the same weight. We don't want the sum to be dominated by a few terms.

Example: Regression through the origin

Let X_i = weight of car i and Y_i = fuel in gallons to go 100 miles.

Model: $Y_i = \theta X_i + \epsilon_i$ where ϵ_i are independent errors with

$$E[\epsilon_i] = 0, \mathrm{Var}(\epsilon_i) = \sigma^2, E[|\epsilon_i|^3] < \infty$$



How to estimate θ from data?

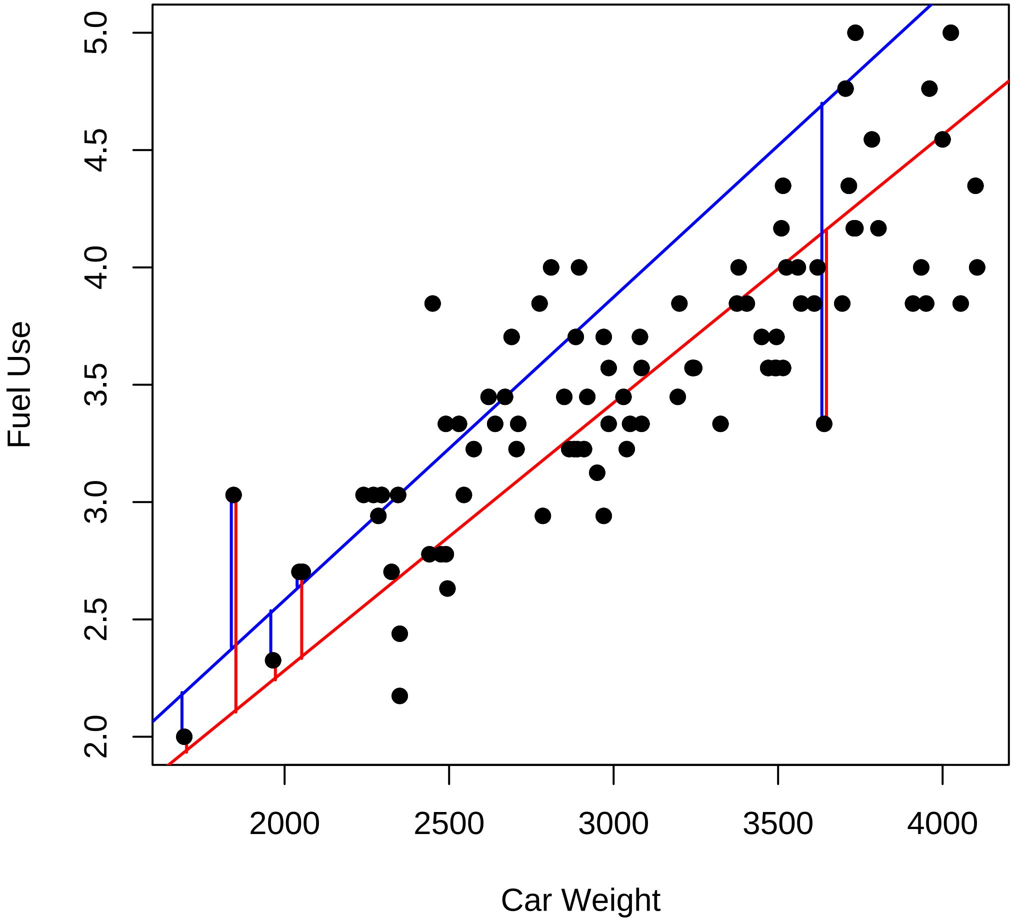
Minimize the least squares criterion

$$SS(\theta) = \sum_{i=1}^{n} (Y_i - \theta X_i)^2$$

which is minimized by

$$\hat{\theta} = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2}$$

What is the distribution of $\hat{\theta} - \theta$?



$$\hat{\theta} = \frac{\sum_{i=1}^{n} X_i(\theta X_i + \epsilon_i)}{\sum_{i=1}^{n} X_i^2} = \theta + \frac{\sum_{i=1}^{n} X_i \epsilon_i}{\sum_{i=1}^{n} X_i^2}$$

Let $Z_i = X_i\epsilon_i$. Thus $E[Z_i] = 0, \mathrm{Var}(Z_i) = X_i^2\sigma^2$. Thus

$$\frac{\sum_{i=1}^n (X_i\epsilon_i - 0)}{\sqrt{\sum_{i=1}^n X_i^2\sigma^2}} \xrightarrow{\mathcal{D}} N(0,1)$$

Note that

$$\frac{\sum_{i=1}^n X_i\epsilon_i}{\sqrt{\sum_{i=1}^n X_i^2\sigma^2}} \times \frac{\sigma}{\sqrt{\sum_{i=1}^n X_i^2}} = (\hat{\theta} - \theta)$$

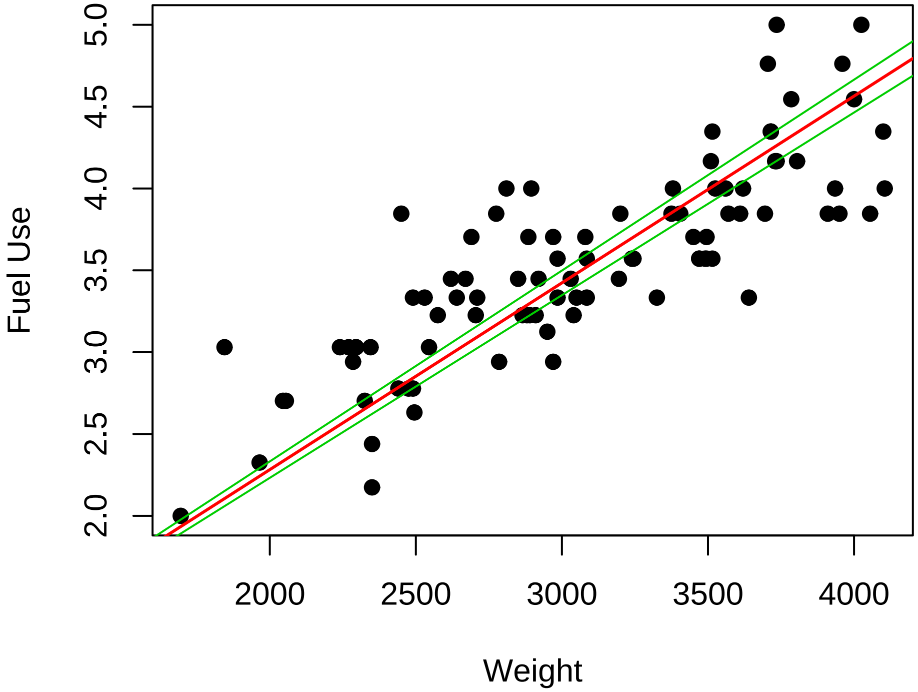
implying

$$(\hat{\theta} - \theta)\sqrt{\sum_{i=1}^n X_i^2} \xrightarrow{\mathcal{D}} N(0,\sigma^2)$$

So

$$\hat{\theta} \overset{approx.}{\sim} N\left(\theta, \frac{\sigma^2}{\sum_{i=1}^n X_i^2}\right)$$

If weight is measured in 100's of pounds, the estimate of θ is $\hat{\theta} = 0.114$ (which implies that each additional 100 pounds of weight appears to add 0.114 gallons to the fuel use on average).



The estimate of σ is $s = 0.3811$. This gives a standard error of

$$\frac{s}{\sqrt{\sum_{i=1}^{93} X_i^2}} = 0.00126$$

which implies we are estimating θ very precisely in this case.

$$\hat{\theta} \overset{approx.}{\sim} N(\theta, 0.00126^2)$$

(Red line: fitted line. Green lines: 95% confidence intervals of the fitted line.)

There are also versions of the CLT that allow the variables to have limited levels of dependency.

They all have the basic form (under different technical conditions)

$$\frac{S_n - E[S_n]}{\mathrm{SD}(S_n)} \xrightarrow{\mathcal{D}} N(0,1) \quad \text{or} \quad \frac{\bar{X}_n - E[\bar{X}_n]}{\mathrm{SD}(\bar{X}_n)} \xrightarrow{\mathcal{D}} N(0,1)$$

which imply

$$S_n \overset{approx.}{\sim} N(E[S_n], \mathrm{Var}(S_n)) \quad \text{or} \quad \bar{X}_n \overset{approx.}{\sim} N(E[\bar{X}_n], \mathrm{Var}(\bar{X}_n))$$

These mathematical results suggest why the normal distribution is so commonly seen with real data.

They say, that when an effect is the sum of a large number of small, roughly equally weighted terms, the effect should be approximately normally distributed.

For example, peoples heights are influenced by (a potentially) large number of genes and by various environmental effects.

Histograms of adult men and women's heights are both well described by normal densities.

Theorem. [Slutsky's Theorems] *Suppose* $X_n \xrightarrow{\mathcal{D}} X$ *and* $Y_n \xrightarrow{P} c$ *(constant). Then*

1. $X_n + Y_n \xrightarrow{\mathcal{D}} X + c$

2. $X_n Y_n \xrightarrow{\mathcal{D}} cX$

3. *If* $c \neq 0$, $\frac{X_n}{Y_n} \xrightarrow{\mathcal{D}} \frac{X_n}{c}$

4. *Let* $f(x, y)$ *be a continuous function. Then* $f(X_n, Y_n) \xrightarrow{\mathcal{D}} f(X, c)$

Example: Suppose $X_1, X_2, \ldots$ are iid with $E[X_i] = \mu, \mathrm{Var}(X_i) = \sigma^2$. What are the distributions of the t-statistics

$$T_n = \frac{\bar{X}_n - \mu}{S_n/\sqrt{n}}$$

as $n \rightarrow \infty$.

As we have seen before

1. By the central limit theorem

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{\mathcal{D}} N(0,1)$$

2. $S_n^2 \xrightarrow{P} \sigma^2$, or $\frac{S_n}{\sigma} \xrightarrow{P} 1$

$$T = \left[\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}\right] \Big/ \frac{S_n}{\sigma} \xrightarrow{\mathcal{D}} \frac{N(0,1)}{1} = N(0,1)$$

This result proves that the t_n distributions converge to the $N(0,1)$ distribution.