LECTURE OUTLINE
Continuos Distributions

Professor Leibon

Math 50

Jan. 19, 2004

Goals

Distribution Functions

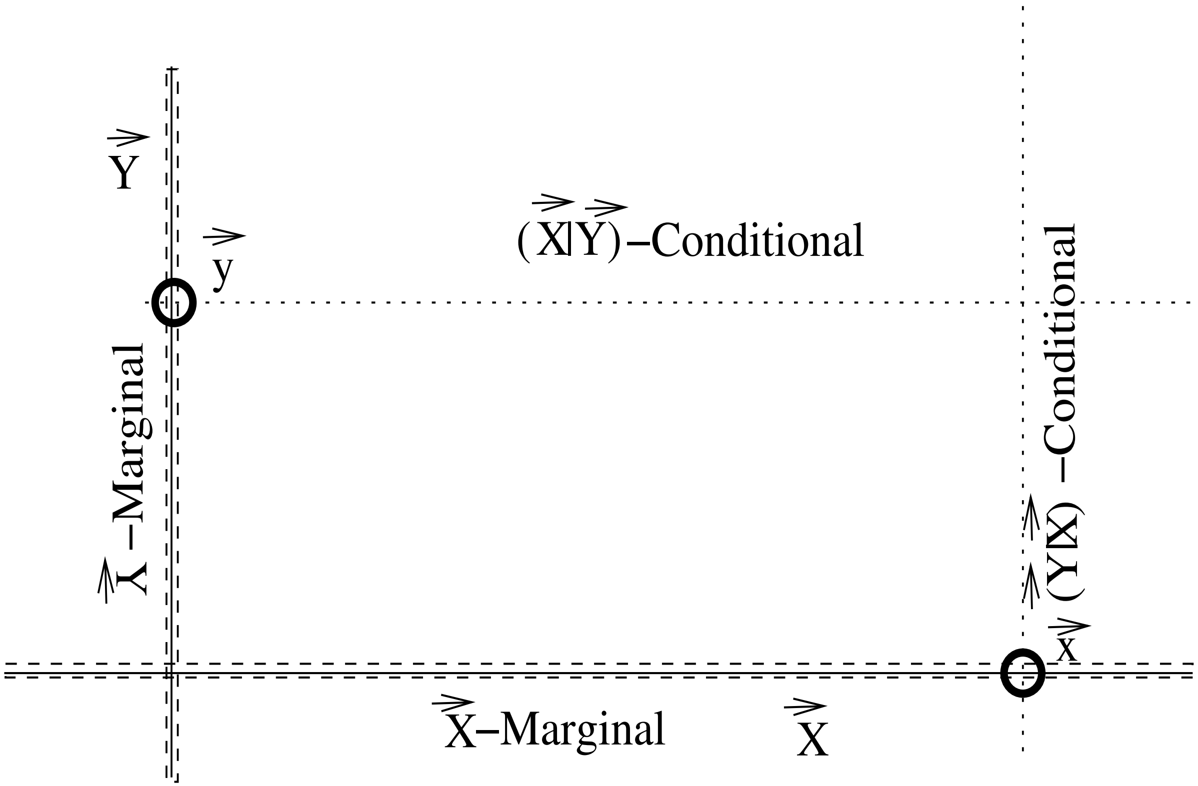
Continuos Distributions

Terminology: A Reference Slide

Function	Abbrv.	Book	Alternate
Probability	pdf or pf	$f(\vec{x})$	$f_{\vec{X}}(\vec{x})$
Distribution	df or cdf	$F(\vec{x})$	$F_{\vec{X}}(\vec{x})$
Quantile	icdf or qf	$F^{-1}(p)$	$Q_X(p)$
$\vec{Y}$-Marginal		$f_2(\vec{y})$	$f_{\vec{Y}}(\vec{y})$
$(\vec{X}\vert\vec{Y})$-Conditional		$g(\vec{x}\vert\vec{y})$	$f_{(\vec{X}\vert\vec{Y}=\vec{y})}(\vec{x})$

For the record: a pdf/pf with some discrete and some continuous coordinates will be called a pdf, and $\int f_{\vec{X}} d\vec{y}$ will mean integrate over the continuous $\vec{y}$-coordinates and sum over the discrete $\vec{y}$-coordinates.

Our Terminology



The Uniform Distribution, $U[0,1]$

$U[0,1]$ is a random variable which returns a real number between 0 and 1 such that if $(a,b],(c,d] \in [0,1]$ and $b-a=d-c$, then $Pr(U[0,1] \in [a,b]) = Pr(U[0,1] \in [a,b])$.

Discuss what this means in "reality". Namely, construct $U[0,1]$ up to some accuracy by flipping a coin.

The Distribution Function

Given a real valued random variable X we let

$$F_X(x) = Pr(X \leq x),$$

and call this function X's *distribution function*, df.

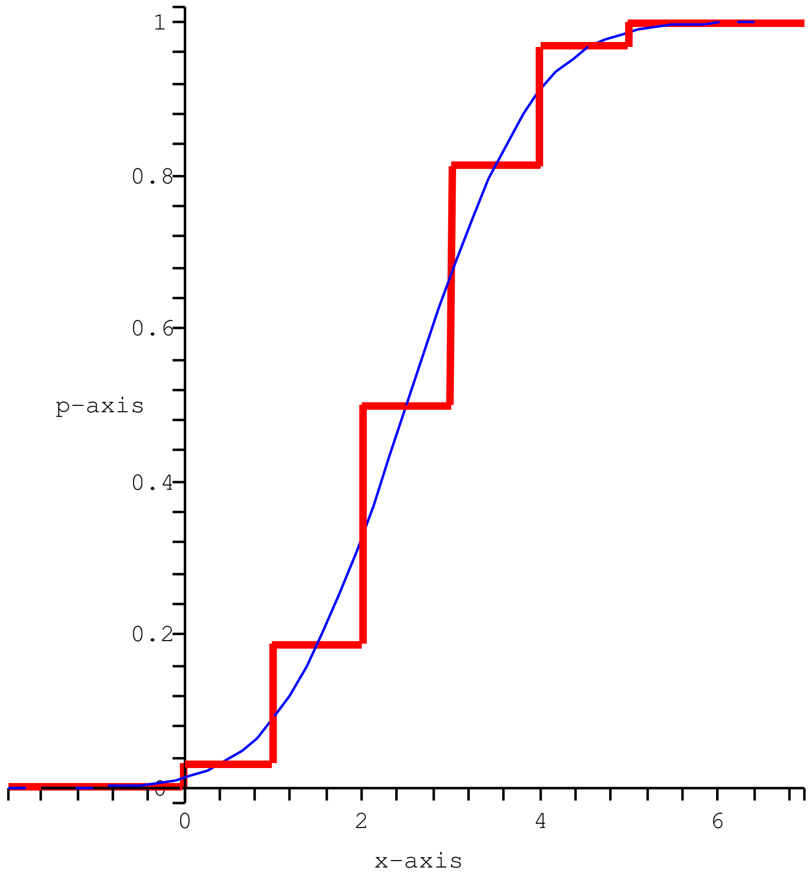
Let

$$Q_X(p) = \min_{p \leq F_X(x)} \{x\} : (0,1) \to \mathbf{R},$$

and call this function X's *quantile function*, qf or icdf.

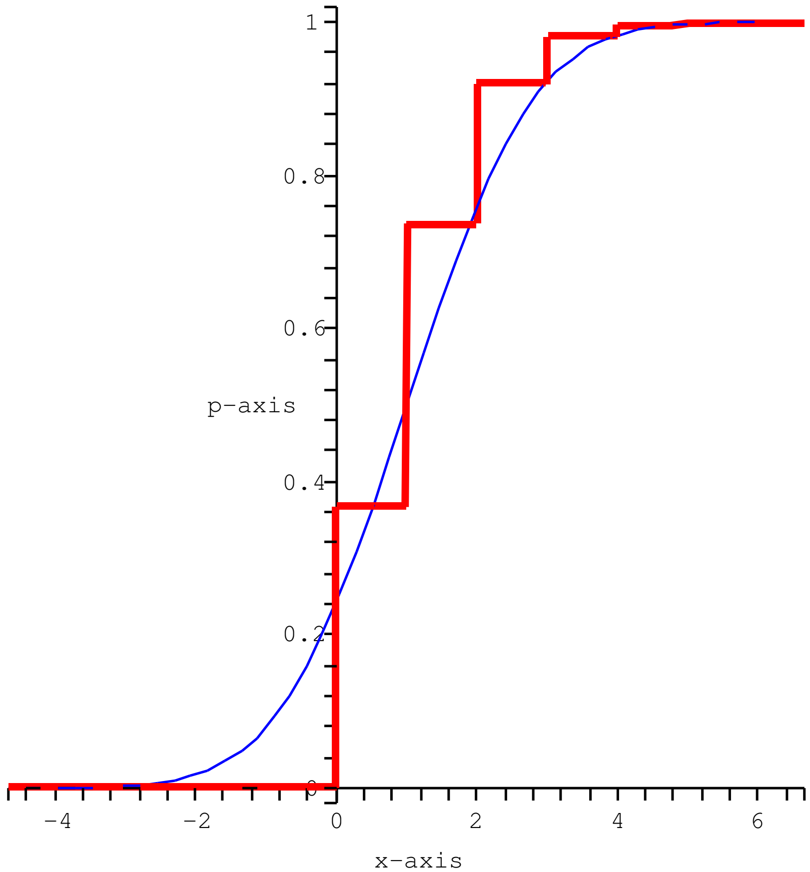
$Bin[5, 0.5]$'s df

Describe how $F_{Bin[5,0.5]}(x)$ and $Q_{Bin[10,0.5]}(p)$ can be found in the following graph.



$Geo[0.5]$'s df

Describe how $F_{Geo[0.5]}(x)$ and $Q_{Geo[0.5]}(p)$ can be found in the following graph.



Summary: df for real valued X

Theorem: $F(x)$ satisfies

- If $x_1 \leq x_2$, then $F(x_1) \leq F(x_2)$
- $\lim_{x \to -\infty} F(x) = 0$
- $\lim_{x \to \infty} F(x) = 1$
- $\lim_{\varepsilon^+ \to 0} F(x + \varepsilon) = F(x)$

if and only if $F(x) = F_X(x)$ for some real valued random variable X. Furthermore, $X = Q_X(U[0, 1])$.

Continuos Random Variable

A real valued random variable X is called *continuos* provided

$$f(x) = \begin{cases} \frac{dF_X}{dx}(x) & \frac{dF_X}{dx} \text{ exist at } x \\ 0 & \frac{dF_X}{dx} \text{ does not exist at } x \end{cases}$$

satisfies $\int_{\infty}^{\infty} f(x)dx = 1$. In this case, we let $f(x) = f_X(x)$ and call f_X one of X's *probability density functions*, pdf.

If X is continuos, then by the Fundamental Theorem of Calculus $P(X \in [a,b]) = P(X \in (a,b)) = \int_a^b f_X(x)dx$.

Example

Let $X = (U[0,1])^2$.

Simulate X with a coin.

To what accuracy do you know you the result of your simulation?

Find F_X.

Is X continuous?

If X is continuous, then find a pdf.

Summary: pdf for real valued X

Theorem: $f(x)$ satisfies

- If $f(x) \geq 0$
- $\int f(x)dx = 1$

then $F_X(x) = \int_{-\infty}^{x} f(x)dx$, is the distribution function of a real valued random variable X. We say $f(x) = f_X(x)$ and call $f_X(x)$ one of X's *probability density functions*.

Comment: We say "one of X's *probability density functions*" since we would not mind if we were to change $f_X(x)$ at say a finite list of points (or on any other *measure zero set*).

Example: The Standard Normal, $N[0,1]$.

Verify

$$f_{N[0,1]}(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$$

is a pdf of a real valued random variable. We call this random variable the *standard normal*, and denote it as $N[0,1]$.

$\mathbf{R}^n$ *Valued Random Variables*

Given real valued random variable $\{X_i\}_{i=1}^N$ we let

$$F_{\vec{X}}(\vec{x}) = Pr(X_1 \leq x_1, \ldots, X_N \leq x_N),$$

and call this function the *joint distribution function* of the $\{X_i\}_{i=1}^N$.

Decide on the correct analog to the results on the summary page for the df of a real valued random variable.

$\mathbf{R}^n$ *valued Random Variables*

If the integral of the following function

$$f(\vec{x}) = \begin{cases} \frac{\partial \cdots \partial F_{\vec{X}}}{\partial x_1 \cdots \partial x_N} & \text{if this derivative exist at } \vec{x} \\ 0 & \text{derivative does not exist at } \vec{x} \end{cases}$$

over $\mathbf{R}^N$ is 1, then we say $f(\vec{x}) = f_{\vec{X}}(\vec{x})$'s and call this function one of X's *joint probability density functions*.

Decide on the correct analog to the results on the summary page for the pdf of a real valued random variable.

Example

Let $X = U[0,1]$ be view as a break of a stick.
After performing this break, break the first half of the stick again and call it $Y = U[0,X]$.
Simulate (X,Y) with a coin.
Find $f_{(X,Y)}$.
What is $Pr(Y > \frac{1}{2})$?

Marginals and Conditionals: Bivariate

Finding f_X from $f_{(X,Y)}$ is called computing the *marginal* of X, i.e.

$$f_X(x) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,y)dy$$

The *conditional* distribution of X given $Y = y$ is given by

$$g_{(X|Y=y)}(x) = \frac{f_{(X,Y)}(x,y)}{f_Y(y)}$$

(when $f_Y(y) > 0$ and arbitrarily otherwise, similarly for $g_{(Y|X=x)}(y)$). Redo the previous example using these concepts.

Bay's Theorem

Notice,

$$g_{(Y|X=x)}(y) = \frac{g_{(X|Y=y)}(x) f_Y(y)}{f_X(x)}.$$

Let's Do Some Statistics!...

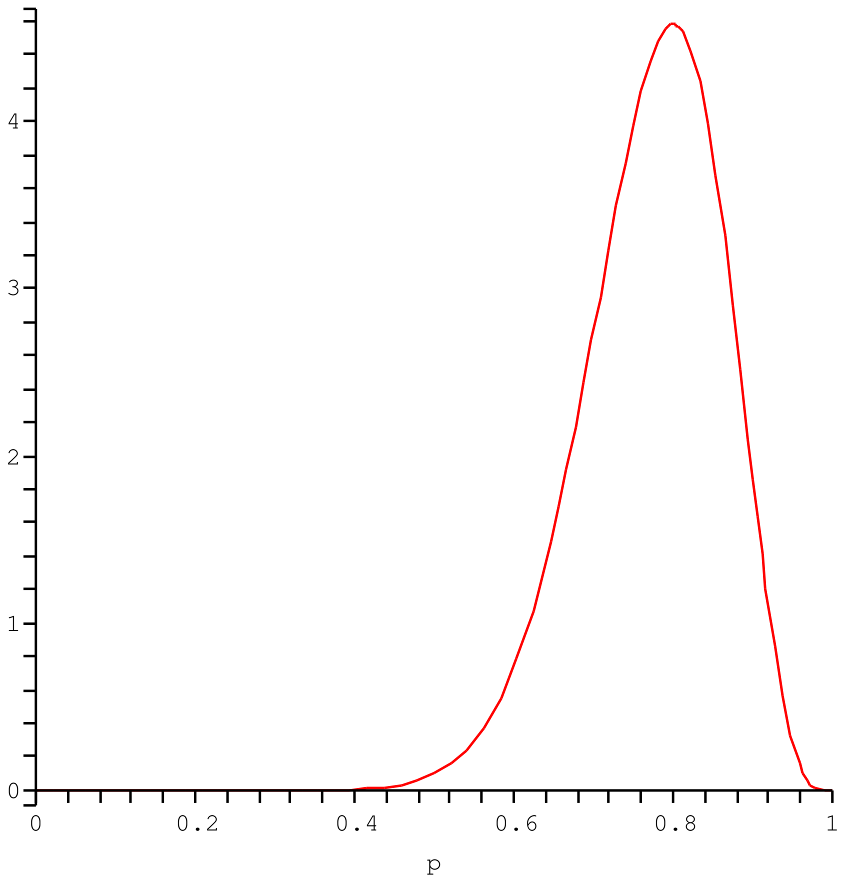
Bay's Theorem

Let's Do Some Statistics!. Suppose that you have a coin which you are going to test if it is loaded. If its near $p = 1/2$ say $[.45, .55]$ you will view as not relevantly loaded. Otherwise you will view it as relevantly loaded.
If it is a VERY strange shaped coin, you MIGHT choose your prior to be $p = U[0, 1]$. Run the success failure experiment once, and let X be the Bernoulli RV which is 1 when a success occurs. Suppose you see a success. **What is p's posterior distribution?**

If you run the experiment a second time and see a success, what is what will be our new view of p's distribution?

Bay's Theorem

I rolled $X := [1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 0, 1]$, and hence found $g(p \mid X) = 101745p^{16}(1-p)^4$.



The Law of Large Numbers, WLLN

Suppose $\{X_i\}$ are i.i.d. with $\mu = E(X_i) < \infty$ and $\sigma = Sd(X_i) < \infty$. Letting

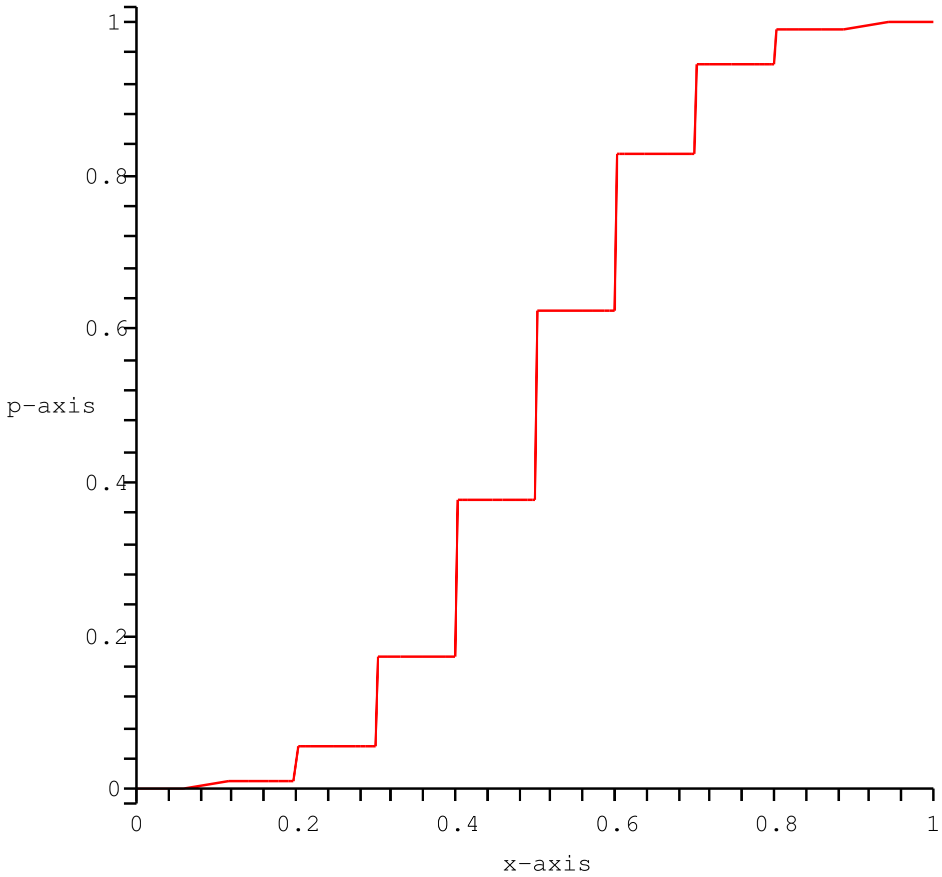
$$A_N = \frac{\sum_{i=1}^{N} X_i}{N},$$

we have that for every $\varepsilon > 0$

$$F_C(\mu + \varepsilon) - F_C(\mu - \varepsilon) \geq 1 - \frac{\sigma^2}{\varepsilon^2} \frac{1}{N}.$$

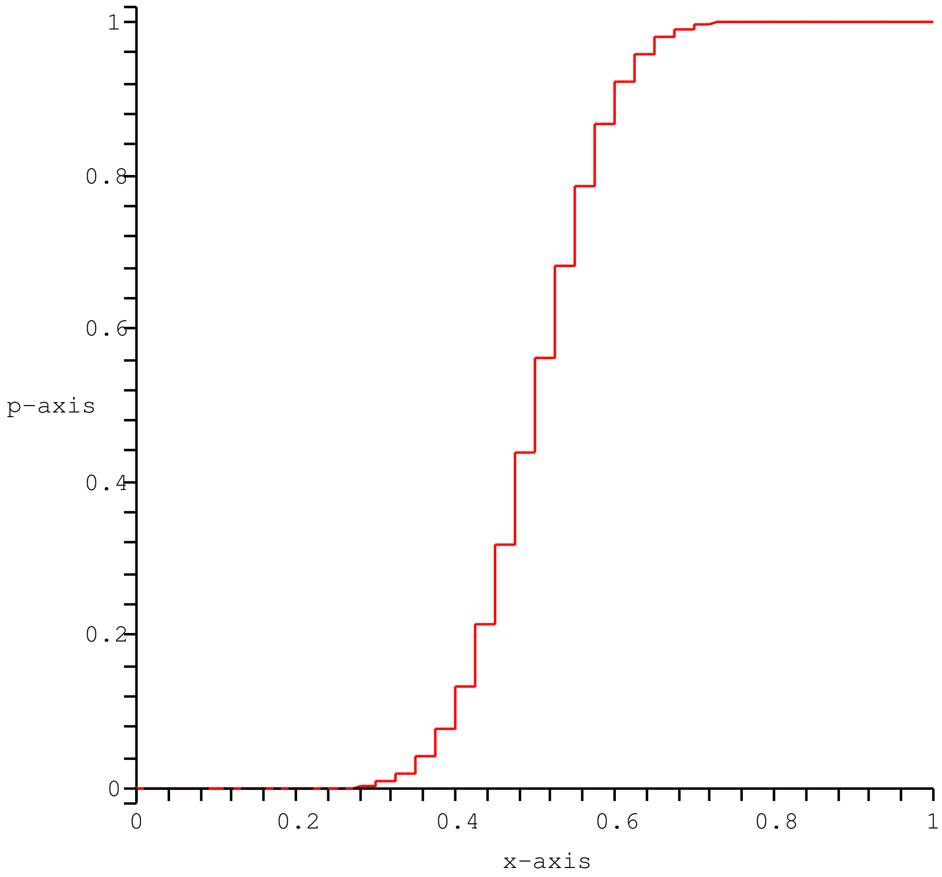
$N = 10.$

Let $X_i = Ber_i[0.5]$, and view A_N verses $F_{N[0,1]}$ for $N = 10$.



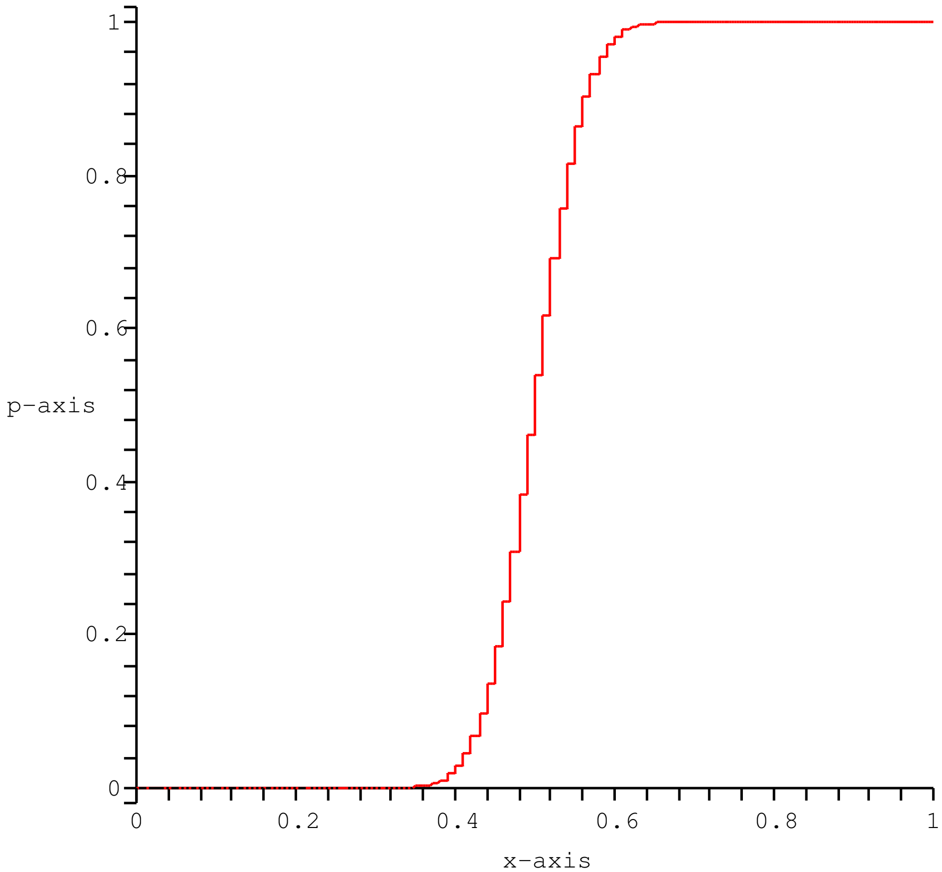
$N = 40.$

Let $X_i = Ber_i[0.5]$, and view A_N verses $F_{N[0,1]}$ for $N = 40$.



$N = 100.$

Let $X_i = Ber_i[0.5]$, and view A_N verses $F_{N[0,1]}$ for $N = 100$.



Central Limit Theorem

Using the notation from the WLLN slide, if $E(|X_i - E(X_i)|^3) = \rho < \infty$ and we let

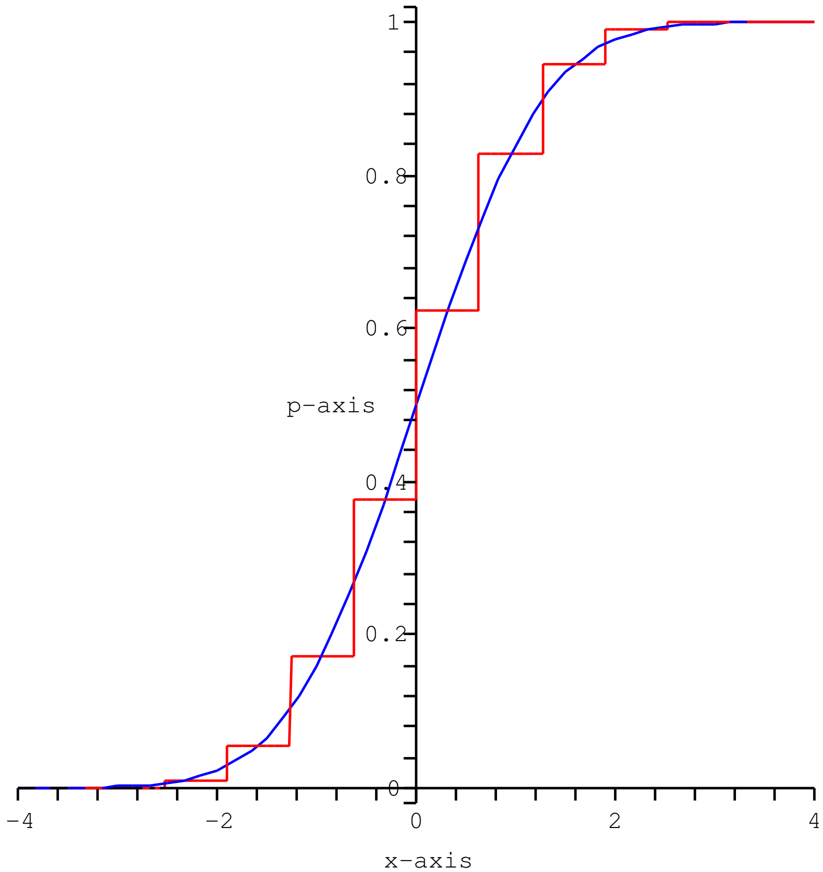
$$S_N^* = \frac{\left(\sum_{i=1}^N X_i\right) - N\mu}{\sigma\sqrt{N}},$$

then for every x

$$\left|F_{S_N^*}(x) - F_{N[0,1]}(x)\right| \leq \frac{\rho}{\sigma^3}\frac{1}{\sqrt{N}}.$$

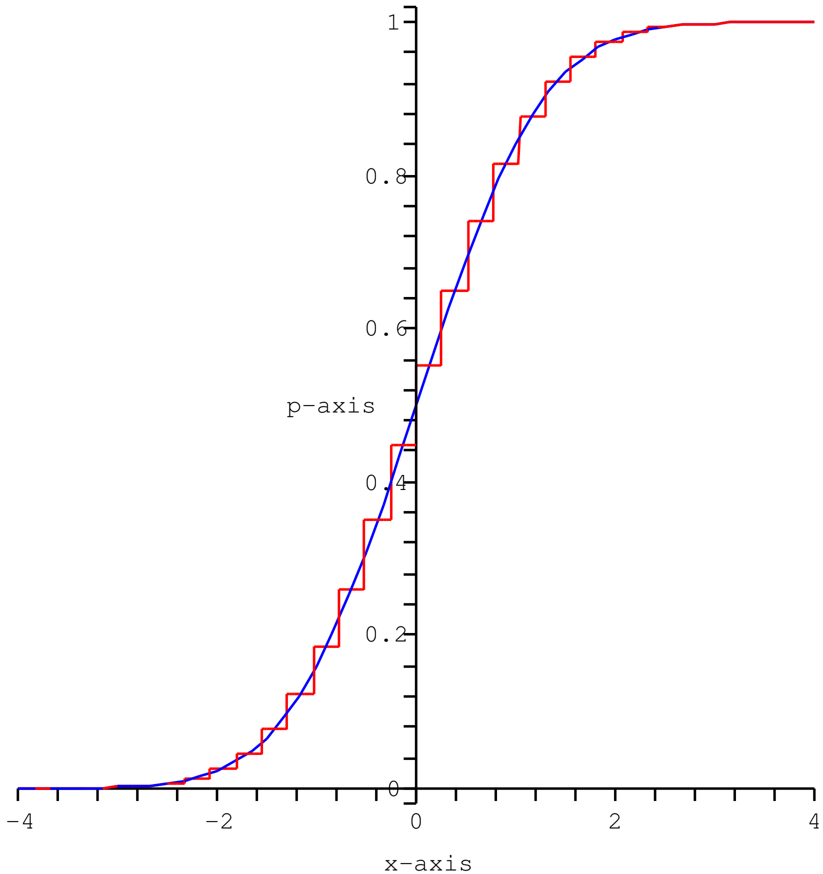
$N = 10.$

Let $X_i = Ber_i[0.5]$, and view S_N^* verses $F_{N[0,1]}$ for $N = 10$.



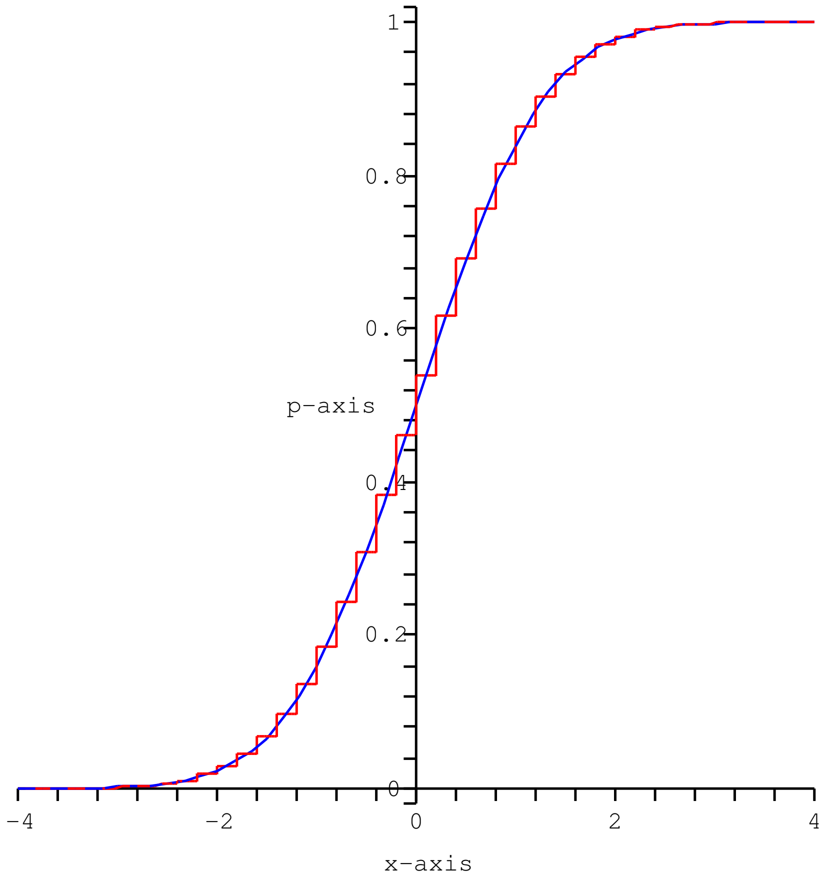
$N = 60.$

Let $X_i = Ber_i[0.5]$, and view S_N^* verses $F_{N[0,1]}$ for $N = 60$.



$N = 100.$

Let $X_i = Ber_i[0.5]$, and view S_N^* verses $F_{N[0,1]}$ for $N = 100$.



$N = 300.$

Let $X_i = Ber_i[0.5]$, and view S^*_N verses $F_{N[0,1]}$ for $N = 300$.