Continuous Probability Functions

Daniel B. Rowe, Ph.D.

Professor
Department of Mathematical and Statistical Sciences

Outline

- Continuous RVs, PDFs, CDFs
- Continuous Expectations
- Continuous Moments
- Continuous Uniform PDFs, CDFs, Matlab
- Continuous Normal PDFs, CDFs, Matlab

Continuous RVs, PDFs, and CDFs

Assume that the continuous random variable (RV) can take on values

$$x \in [a,b]$$

then, the probability distribution function (PDF) is given by

$f(x \mid \theta)$ defined for $x \in [a,b]$

where x can be defined within an infinite interval

and θ are any parameters that the PDF depends on.

Continuous RVs, PDFs, and CDFs

Further, the cumulative distribution function (CDF) is given by

$$F(x\,|\,\theta)=\int_{t=-\infty}^{x} f(t\,|\,\theta)dt$$

Additionally, any PDF must satisfy

1) $0 \le f(x\,|\,\theta)$

2) $\int_{x} f(x\,|\,\theta)=1$.

Continuous Expectation

Given an arbitrary continuous probability distribution

$f(x|\theta)$, we want to compute quantitative population

summaries of it such as:

population mean, $\mu = \int_{x=-\infty}^{\infty} x f(x \mid \theta) dx$

population variance, $\sigma^2 = \int_{x=-\infty}^{\infty} (x-\mu)^2 f(x \mid \theta) dx$

population standard deviation, $\sigma = \sqrt{\sigma^2}$

Continuous Summaries

Given an arbitrary continuous probability distribution

$f(x|\theta)$, we want to compute quantitative population

summaries of it such as:

population median $\tilde{x}$, $\int_{x=-\infty}^{\tilde{x}} f(x|\theta)dx = \frac{1}{2}$

population mode $\hat{x}$, $\left.\frac{\partial}{\partial x} f(x|\theta)\right|_{\hat{x}} = 0$

Provided f is differentiable.
Max if 2nd deriivitive neg at point.
Check boundary points for max.

Continuous Expectation

These population numerical summaries are found by expectation

$$E[g(X)\,|\,\theta] = \int_{x=-\infty}^{\infty} g(x)\, f(x\,|\,\theta) dx$$

The mean is

$$\mu = E(X\,|\,\theta)$$

and the variance is

$$\sigma^2 = E[(X-\mu)^2\,|\,\theta] \quad .$$

Continuous Expectation

Linearity Property:

$$E[\sum_{j=1}^{n} a_j g_j(X) \mid \theta] = \sum_{j=1}^{n} a_j E[g_j(x) \mid \theta]$$

i.e. $E[a_1 g_1(X) + a_2 g_2(X) \mid \theta] = a_1 E[g_1(x) \mid \theta] + a_2 E[g_2(x) \mid \theta]$

$$\begin{aligned} E[a_1 g_1(X) + a_2 g_2(X) \mid \theta] &= \int_{x=-\infty}^{\infty} [a_1 g_1(x) + a_2 g_2(x)] f(x \mid \theta) dx \\ &= a_1 \int_{x=-\infty}^{\infty} g_1(x) f(x \mid \theta) dx + a_2 \int_{x=-\infty}^{\infty} g_2(x) f(x \mid \theta) dx \\ &= a_1 E[g_1(X) \mid \theta] + a_2 E[g_2(X) \mid \theta] \end{aligned}$$

Continuous Moments

For each integer n, the n^{th} moment of X, μ_1', is

$$\mu_n' = E(X^n \mid \theta) \quad .$$

The n^{th} central moment of X, μ_n, is

$$\mu_n = E[(X - \mu)^n \mid \theta]$$

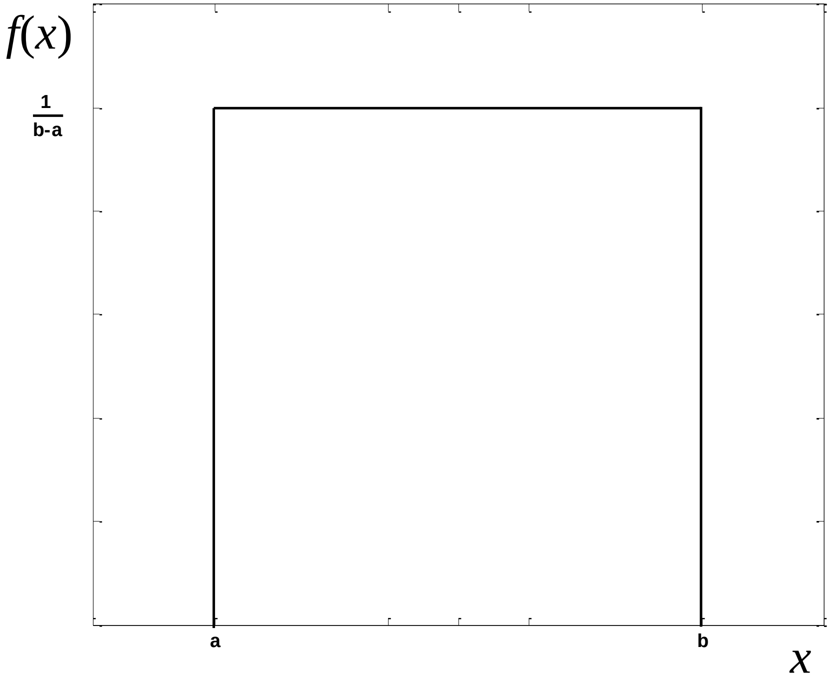
where $\mu_1' = E(X \mid \theta) \quad .$

Continuous Distributions

Uniform:

A random variable x has a continuous uniform distribution, x ~uniform(a,b) if

$$f(x \mid a,b) = \frac{1}{b-a} \;, \quad x \in [a,b]$$



where, $a,b \in \mathbb{R}$, $a < b$.

Continuous Distributions

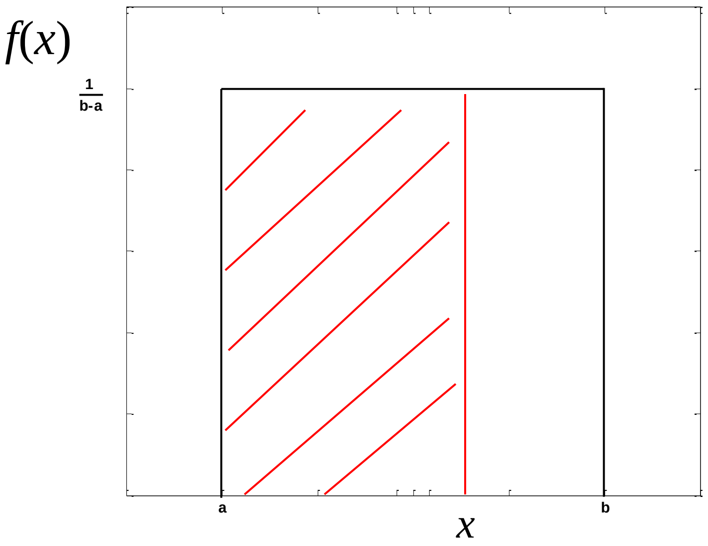
Uniform:

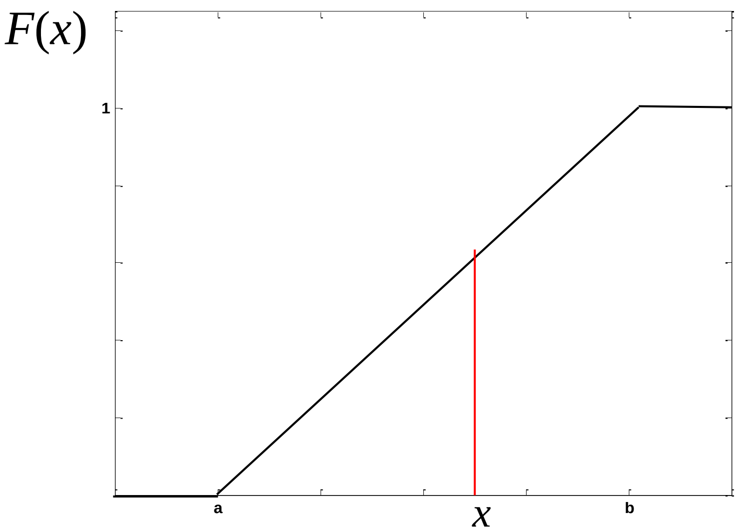
The CDF of the continuous uniform distribution is

$$F(x|\theta) = \int_{t=-\infty}^{x} f(t|\theta)dt$$

$$F(x|a,b) = \int_{t=a}^{x} \frac{1}{b-a}dt$$

$$= \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & x \in [a,b] \\ 1 & x > b \end{cases}$$



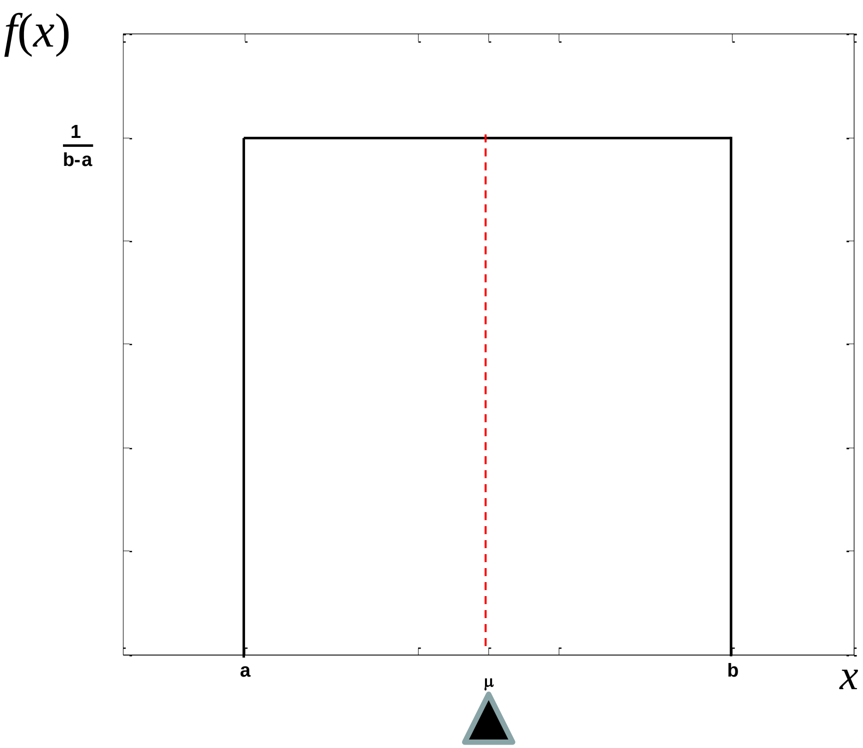


Continuous Distributions

Uniform:

It can be shown that

$$\mu = \int_x x f(x|\theta)dx = \int_{x=a}^{b} x \frac{1}{b-a} dx = \frac{b+a}{2}$$



Continuous Distributions

Uniform:

It can be shown that

median

$$\int_{x=a}^{\tilde{x}} f(x|\theta)dx = \frac{1}{2}$$

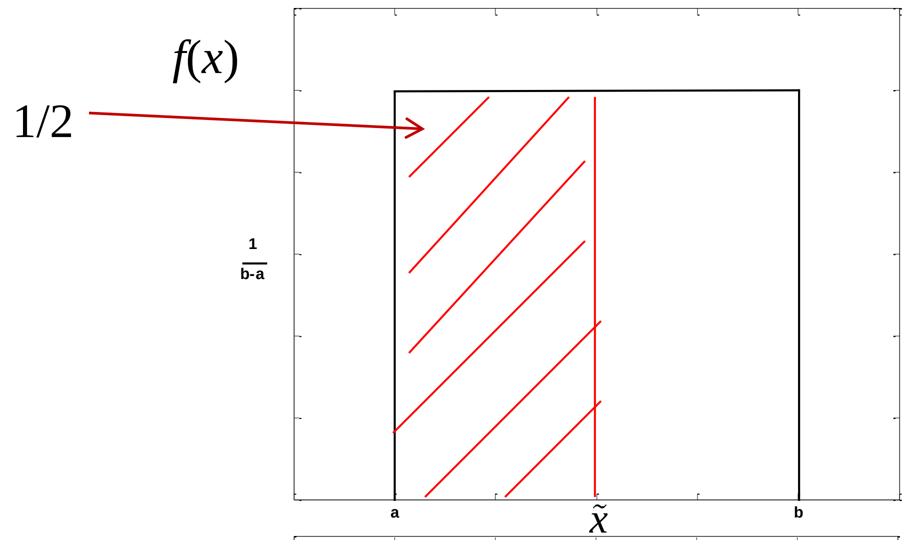
$$\tilde{x} = \frac{b+a}{2}$$

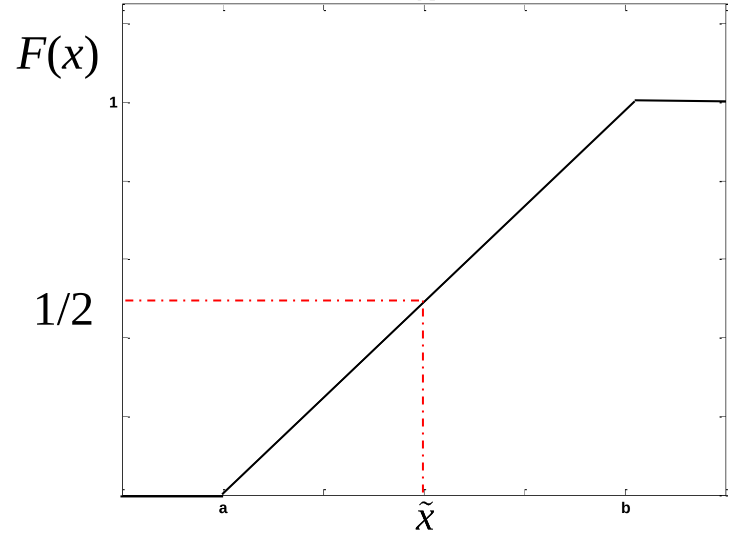
mode

$$\frac{\partial}{\partial x} f(x|\theta)\Big|_{\hat{x}} = 0$$

$$\hat{x} = \text{any value} \in [a,b]$$

flat distribution





Continuous Distributions

Uniform:

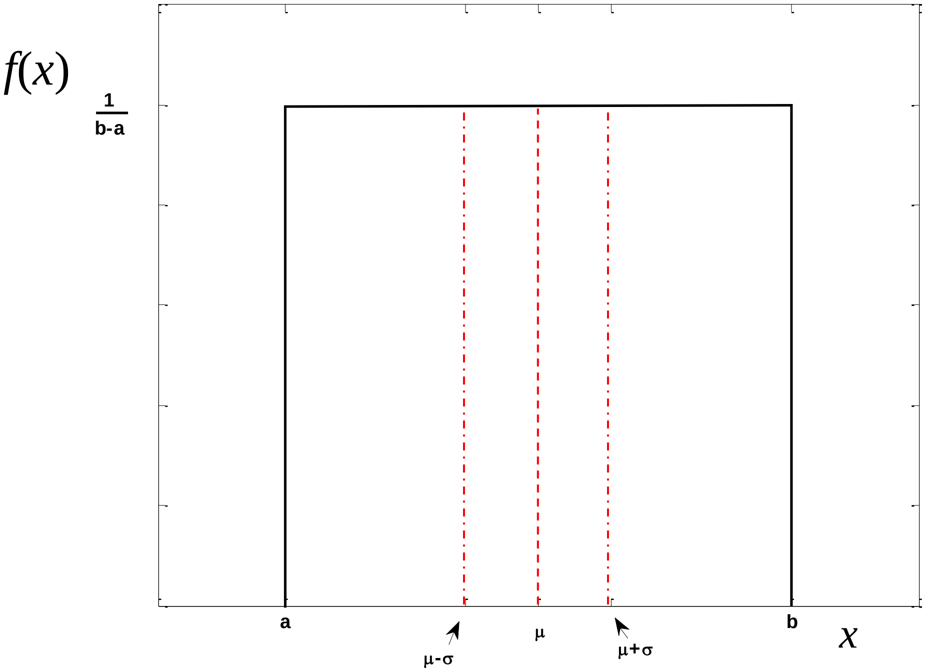
that

$$\sigma^2 = \int_x (x-\mu)^2 f(x|\theta)dx$$

$$= \int_{x=a}^{b} (x-\mu)^2 \frac{1}{b-a} dx$$

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

$$\sigma = \frac{(b-a)}{\sqrt{12}}$$

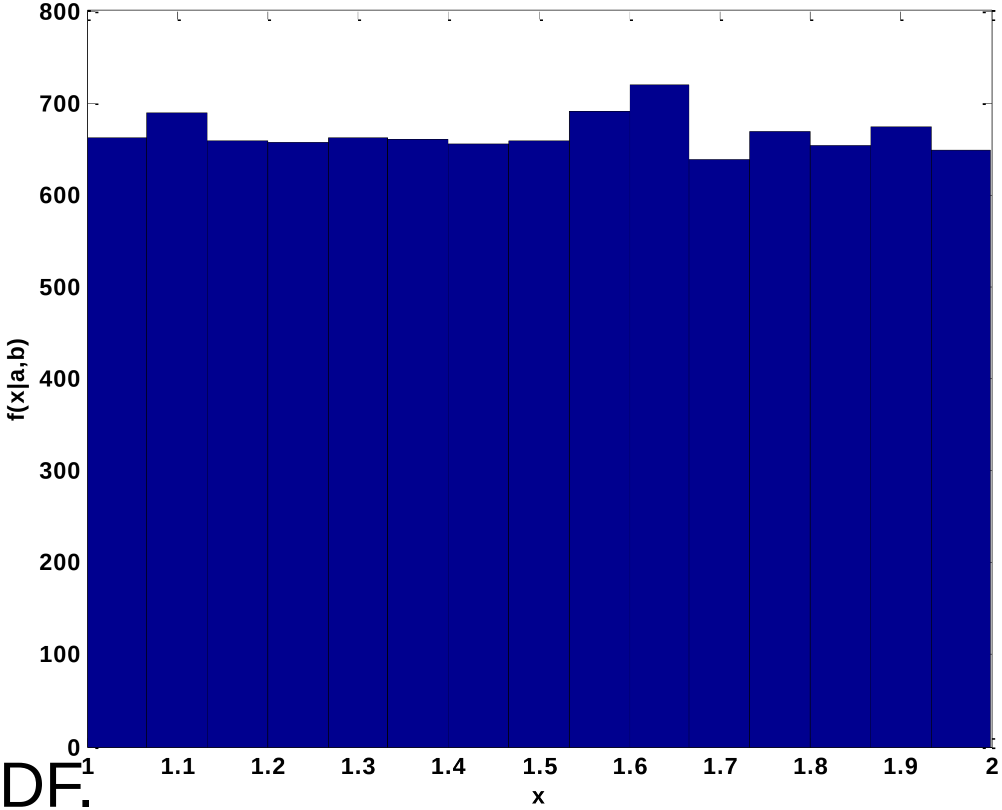


Continuous Distributions

Uniform:

$$f(x \mid a,b) = \frac{1}{b-a}$$

```
a=1;,b=2;,num=10^4;
x=a+(b-a)*rand(num,1);
mean(x)
var(x)
hist(x,15)
```



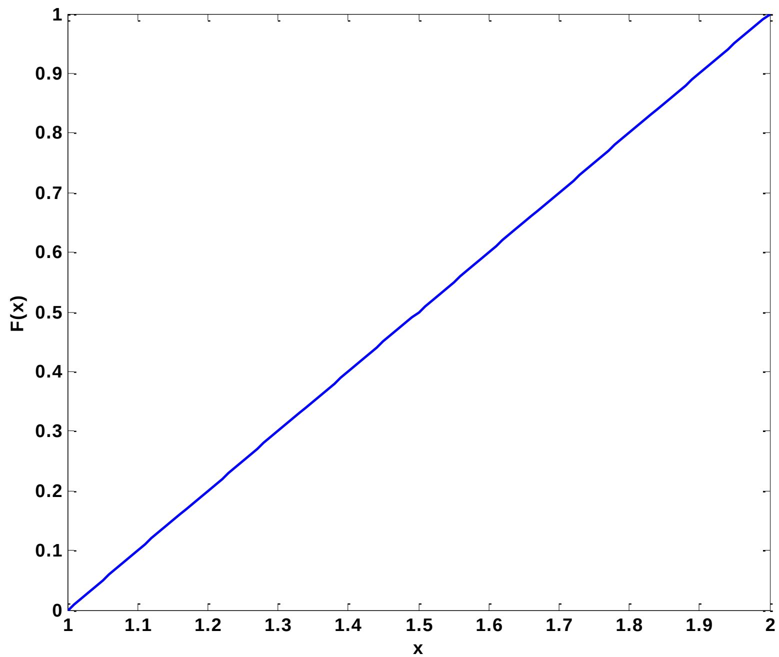
	True	Simulated
μ	1.5	1.4996
σ^2	0.0833	0.0829

Can also find and plot ECDF.

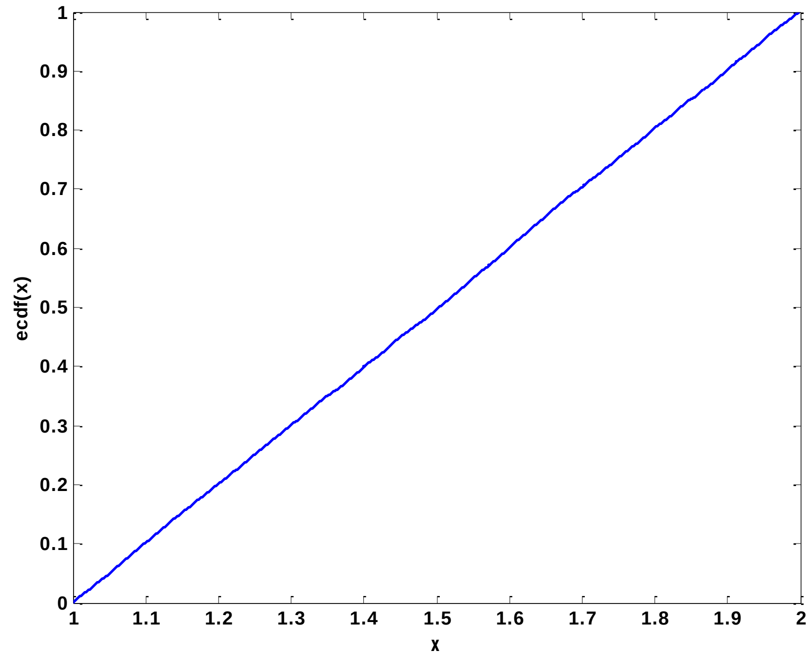
Continuous Distributions

Uniform:

$$f(x|a,b)=\frac{1}{b-a}$$




```
a=1;,b=2;,
y =cdf('unif',(1:.01:2),a,b)
plot((1:.01:2),y)
```




```
[F,xx]=ecdf(x);
stairs(xx,F,'LineWidth',2)
```

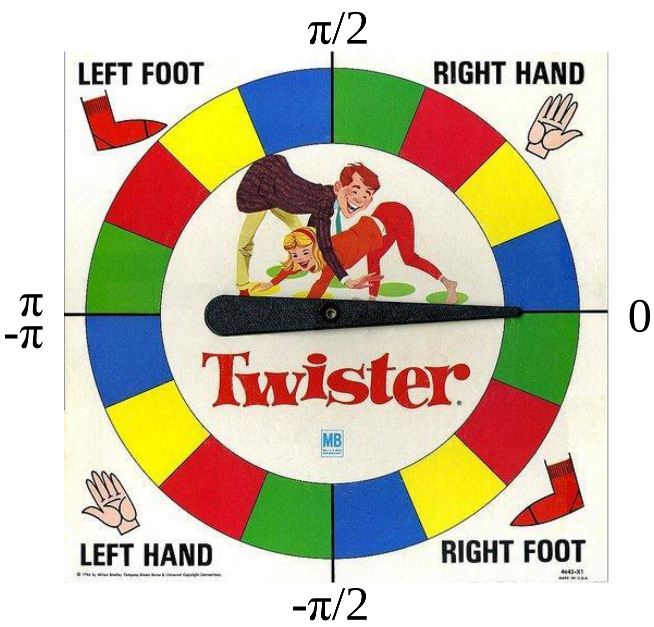
Continuous Distributions

Uniform:

$$f(x \mid a,b) = \frac{1}{b-a}$$

We are playing a board game. See spinner below.

Assume that resulting angle x from a spin is uniformly distributed. i.e. Let x~uniform(a=-π,b=π).



Find $P(red)$.

FYI there are circular distributions!

Continuous Distributions

Normal:

A random variable x has a continuous normal distribution, x ~normal(μ,σ^2) if

$$f(x\,|\,\mu,\sigma^2)=\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}\quad,\quad x\in\mathbb{R}$$

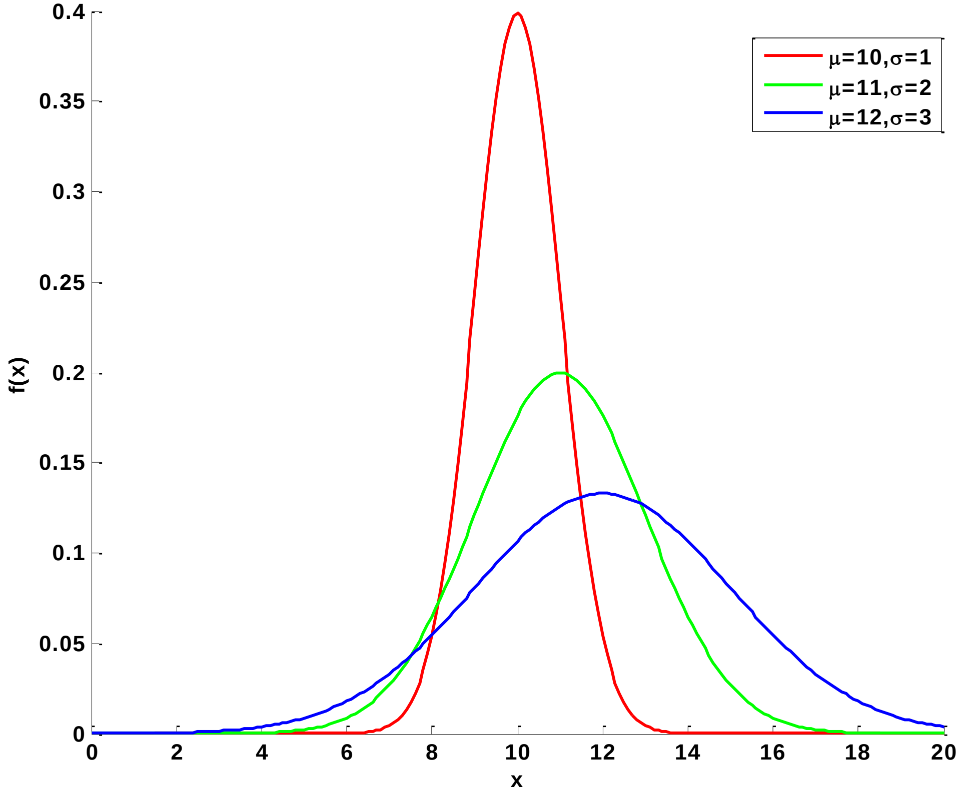
where, $-\infty<\mu<\infty$, and $0<\sigma$.

Continuous Distributions

Normal:

$$f(x\,|\,\mu,\sigma^2)=\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$$

```
x=(0:.1:20)';
mu=[10,11,12];, sigma=[1,2,3];
figure(1)
hold on
for count=1:length(sigma)
    y = normpdf(x,mu(count),sigma(count));
    if count==1
        plot(x,y,'r','LineWidth',2)
    elseif count==2
        plot(x,y,'g','LineWidth',2)
    elseif count==3
        plot(x,y,'b','LineWidth',2)
    end
end
xlim([0 20])
```



Continuous Distributions

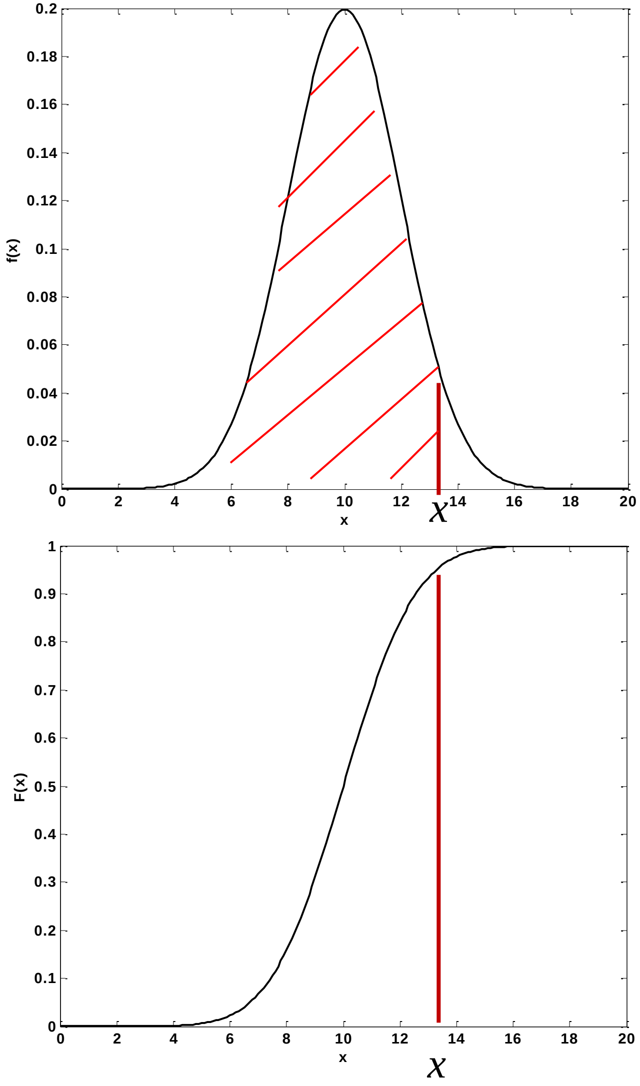
Normal:

The CDF of the continuous normal distribution is

$$F(x|\theta) = \int_{t=-\infty}^{x} f(t|\theta)dt$$

$$F(x|\mu,\sigma^2) = \int_{t=-\infty}^{x} \frac{e^{-\frac{(t-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}dt$$

No closed form analytic solution.

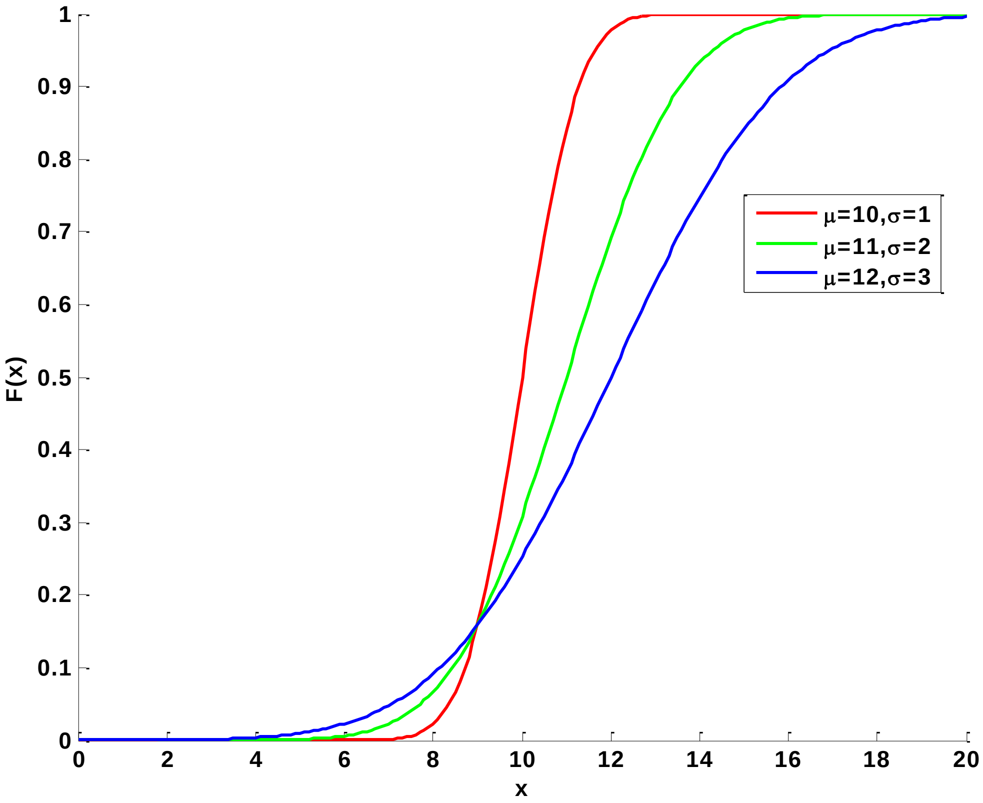


Continuous Distributions

Normal:

$$f(x\,|\,\mu,\sigma^2)=\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$$

```
x=(0:.1:20)';
mu=[10,11,12];, sigma=[1,2,3];
figure(1)
hold on
for count=1:length(sigma)
    y = normcdf(x,mu(count),sigma(count));
    if count==1
        plot(x,y,'r','LineWidth',2)
    elseif count==2
        plot(x,y,'g','LineWidth',2)
    elseif count==3
        plot(x,y,'b','LineWidth',2)
    end
end
xlim([0 20])
```



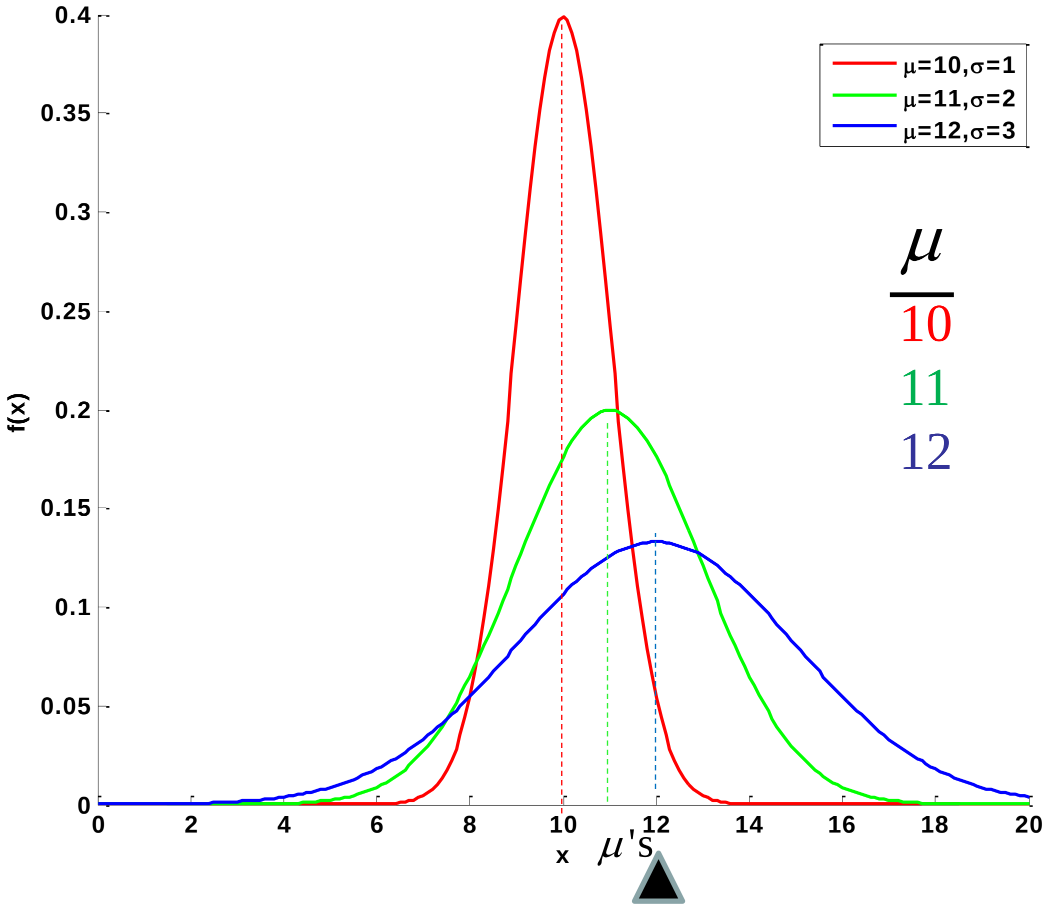
Continuous Distributions

Normal:

It can be shown that

$$\mu = \int_x x f(x|\theta)dx$$

$$= \int_{x=-\infty}^{\infty} x \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}dx$$



Continuous Distributions

Normal:

It can be shown that

1/2

median

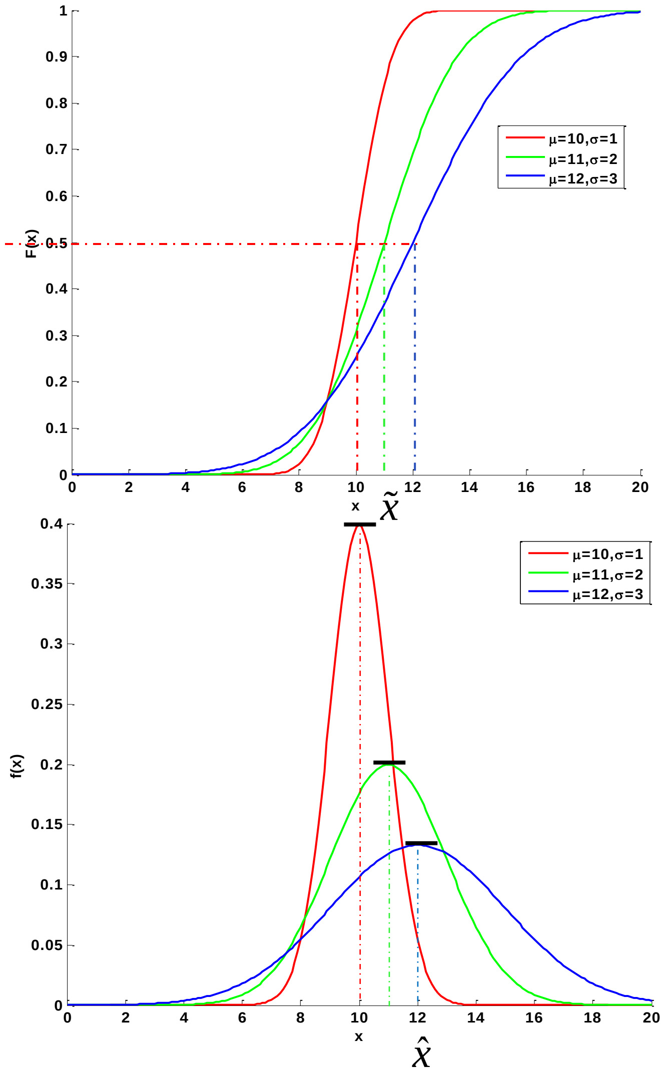
$$\int_{x=-\infty}^{\tilde{x}} f(x\,|\,\theta)dx = \frac{1}{2}$$

$$\tilde{x} = \mu$$

mode

$$\frac{\partial}{\partial x} f(x\,|\,\theta)\Big|_{\hat{x}} = 0$$

$$\hat{x} = \mu$$



$\tilde{x}$
10
11
12

$\hat{x}$
10
11
12

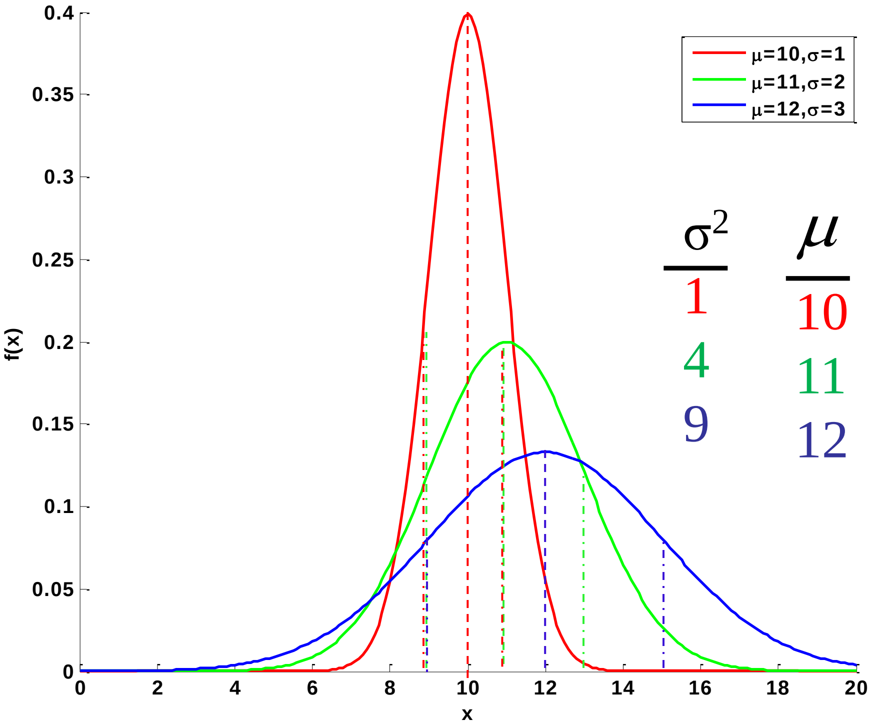
Continuous Distributions

Normal:

that

$$\sigma^2 = \int_x (x-\mu)^2 f(x|\theta)dx$$

$$= \int_{x=-\infty}^{\infty} (x-\mu)^2 \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx$$

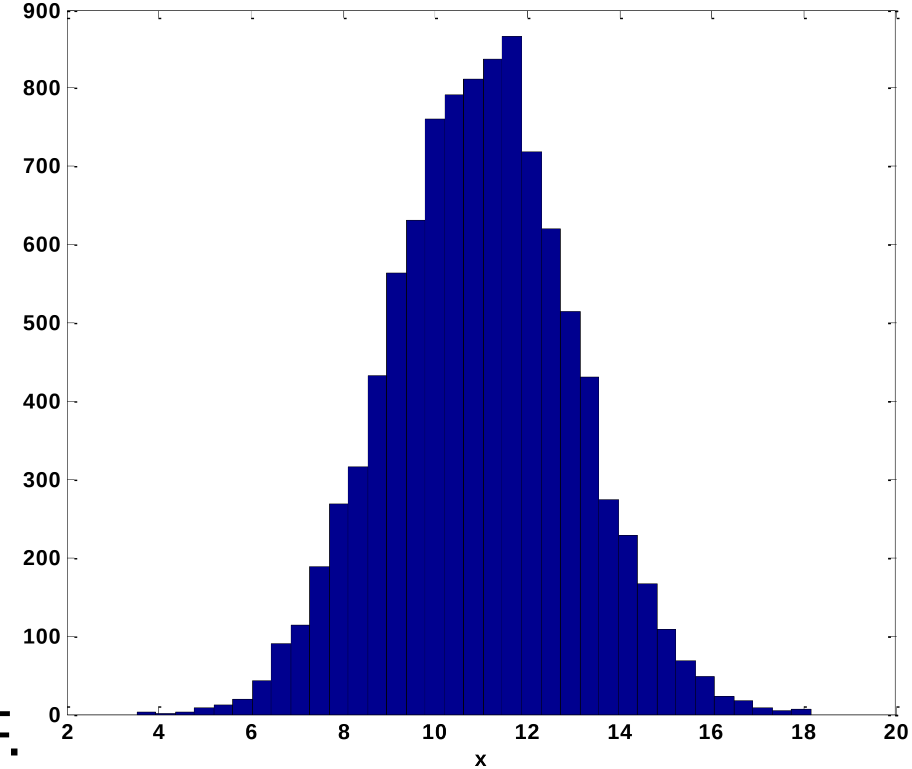


Continuous Distributions

Normal:

$$f(x\,|\,\mu,\sigma^2)=\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$$

```
mu=11;,sigma=2;,num=10^4;
x=normrnd(mu,sigma,num,1);
mean(x)
var(x)
hist(x,35)
```



	True	Simulated
μ	11	11.0033
σ^2	4	3.9321

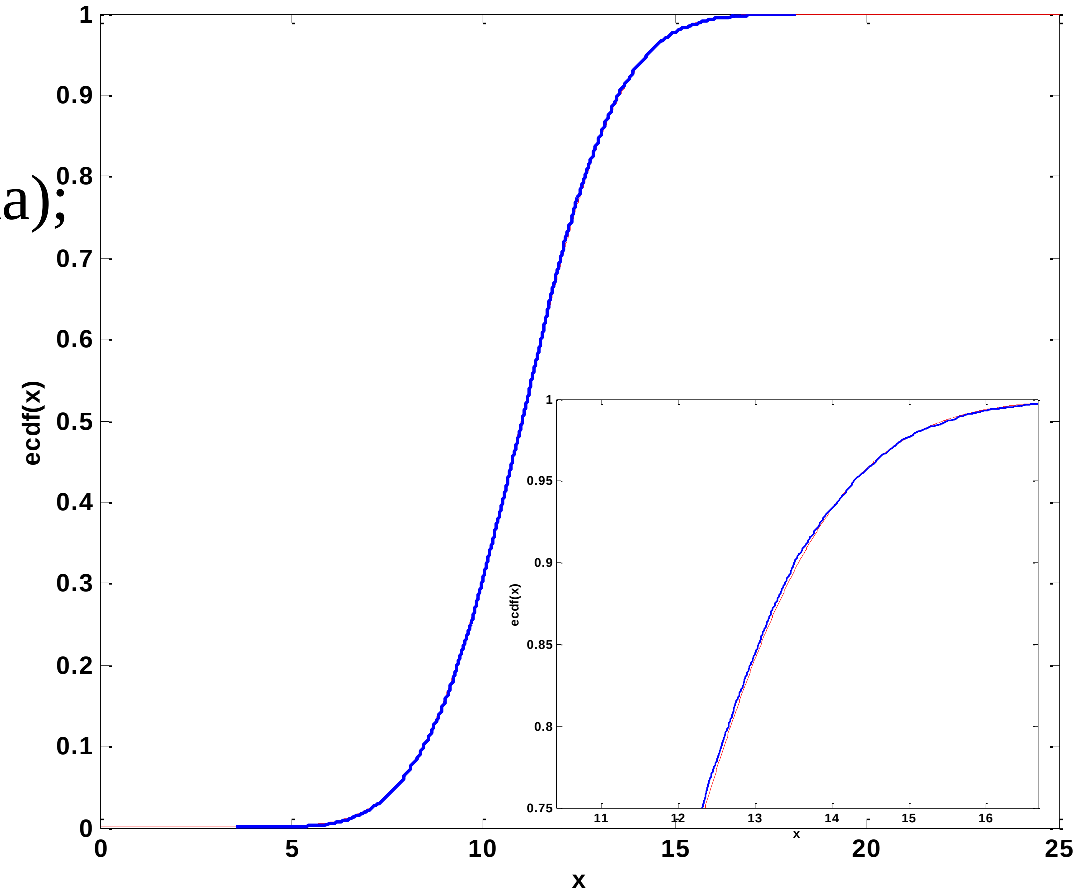
Can also find and plot ECDF.

Continuous Distributions

Normal:

$$f(x\mid\mu,\sigma^2)=\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}$$

```
mu=11;,sigma=2;
y=normcdf((0:.01:25),mu,sigma);
plot((0:.01:25),y,'r')
hold on
[F,xx]=ecdf(x);
stairs(xx,F,'LineWidth',2)
```



Homework 3:

1) Show analytically that $P(\mu-\sigma \leq x \leq \mu+\sigma)=1/\sqrt{3}$ for a uniform distribution. First find μ and σ.

2) Numerically integrate (rectangles) the uniform (0,1) PDF from $\mu-\sigma$ to $\mu+\sigma$ with $\mu=1/2$ and $\sigma=1/\sqrt{12}$.

3) Generate $n=10^6$ random $U(0,1)$ observations. What fraction are within $1/2-1/\sqrt{12}$ to $1/2+1/\sqrt{12}$.

4) Make an empirical CDF from the values in 3).

(Note that this is a Monte Carlo technique for determining percentiles and thus p-values.)

Homework 3:

5) Analytically show that the normal pdf integrates to 1.

6) Analytically show that the mean of normal pdf is μ.

7) Numerically integrate the normal pdf to find the 50th and 99th percentiles for $\mu=11$, $\sigma=2$.
i.e. find x_0 such that $P(x \leq x_0)=.50$ and $P(x \leq x_0)=.99$

8) Generate 10^6 normal $\mu=11$, $\sigma=2$ random variables. Empirically determine the 50th and 99th percentiles. i.e. Find the $.50*10^6$ and $.99*10^6$ largest values x_0. such that $P(x \leq x_0)=.50$ and $P(x \leq x_0)=.99$. Compare to 5).

Homework 3:

9) Make a histogram of the random variables in 8) with at least 50 bins.

10) Make an empirical CDF from the values in 8).

(Note that this is a Monte Carlo technique for determining percentiles and thus p-values.)