Continuous Probability Density Functions

Dr. Daniel B. Rowe
Professor of Computational Statistics
Department of Mathematical and Statistical Sciences
Marquette University

Outline

Properties and Summaries

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

$x \in (a,b)$

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

$f(x|\theta)$ 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.

Properties and Summaries

The accumulation of probability from left to right

is called the cumulative distribution function (PDF)

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

and

1) $0 \leq f(x|\theta)$

2) $\int_x f(x|\theta) = 1$.

Properties and Summaries

Given an arbitrary continuous probability density

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

summaries of it such as:

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

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

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

Properties and Summaries

Given an arbitrary continuous probability density

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

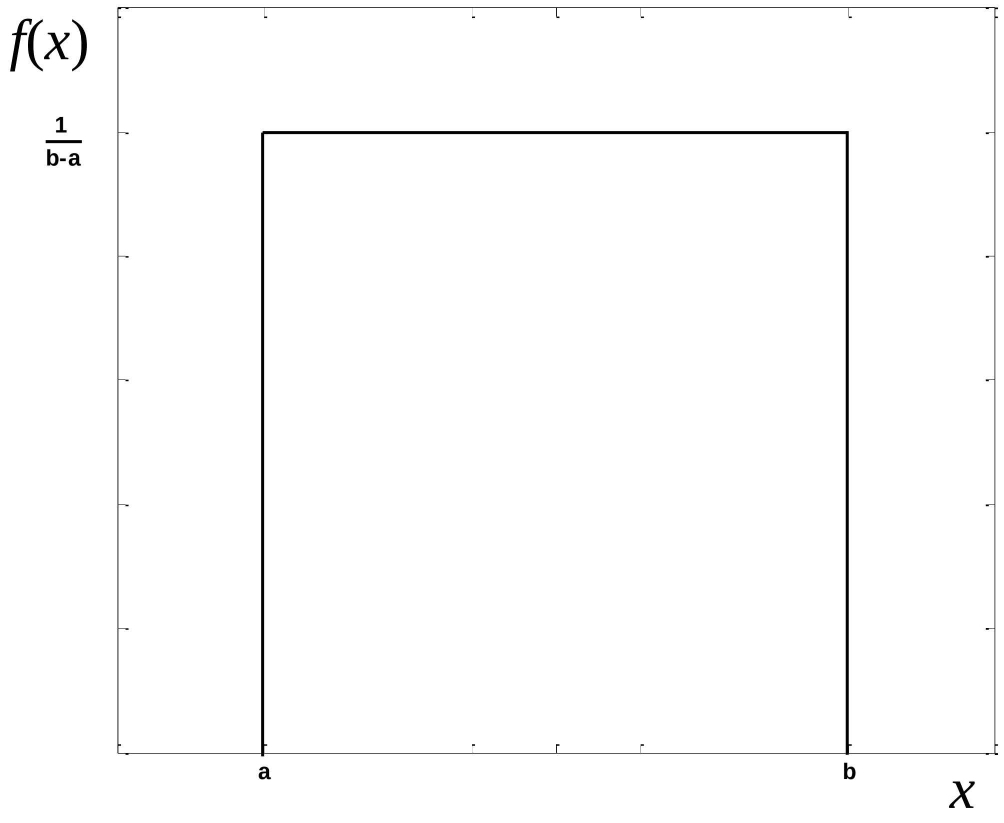
Provided f is differentiable.
Max if 2nd deriivitive neg at point.
Check boundary points for max.
May need to be numerically calculated.

Uniform and Beta PDF

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

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

where, $a<x<b$, $a<b$.



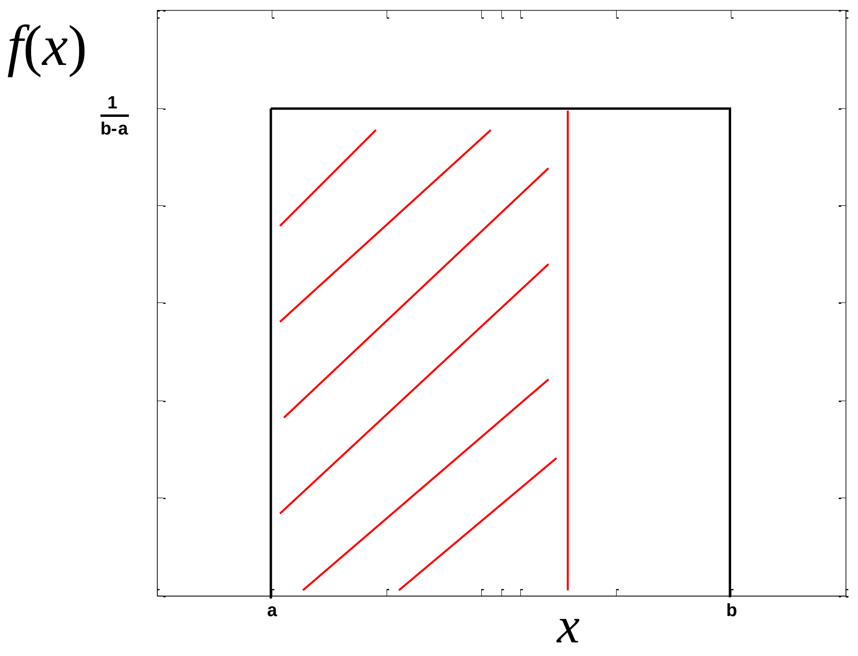
Uniform and Beta PDF

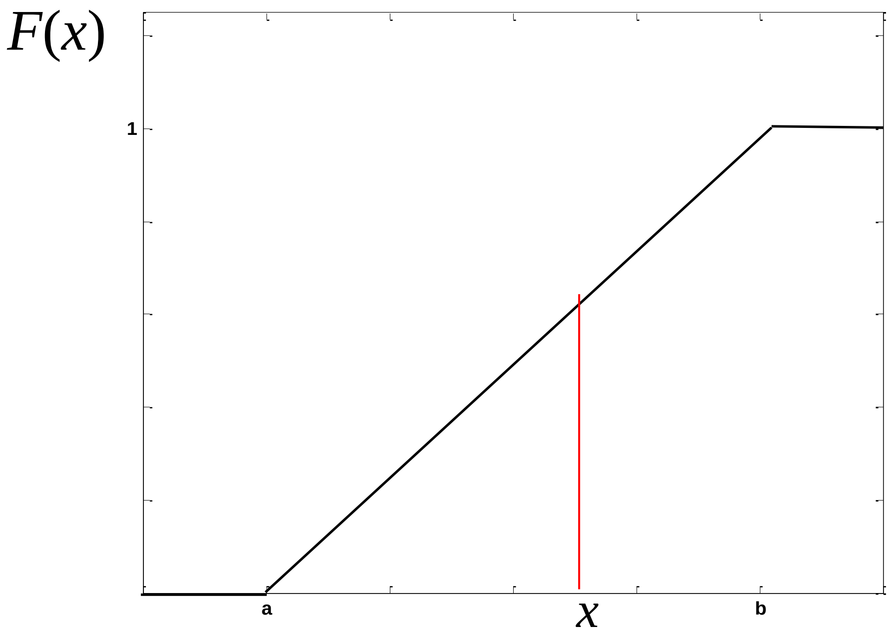
The CDF of the continuous uniform distribution is

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

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

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

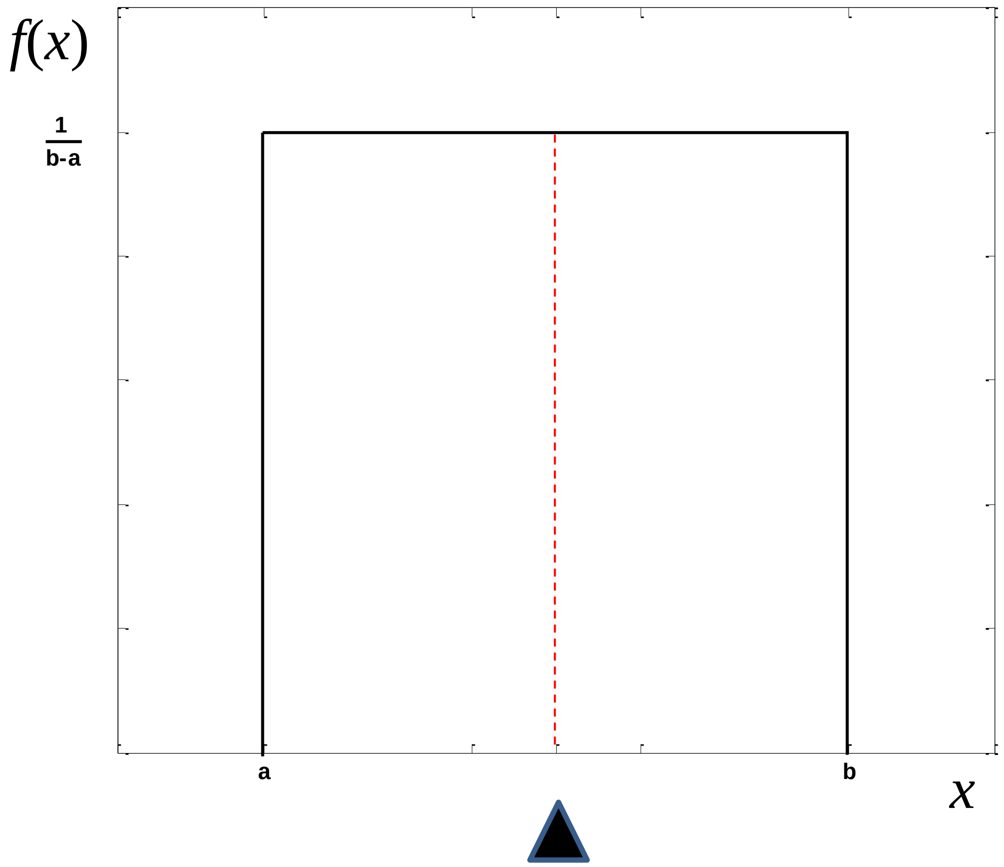




Uniform and Beta PDF

It can be shown that

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



Uniform and Beta PDF

It can be shown that

median

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

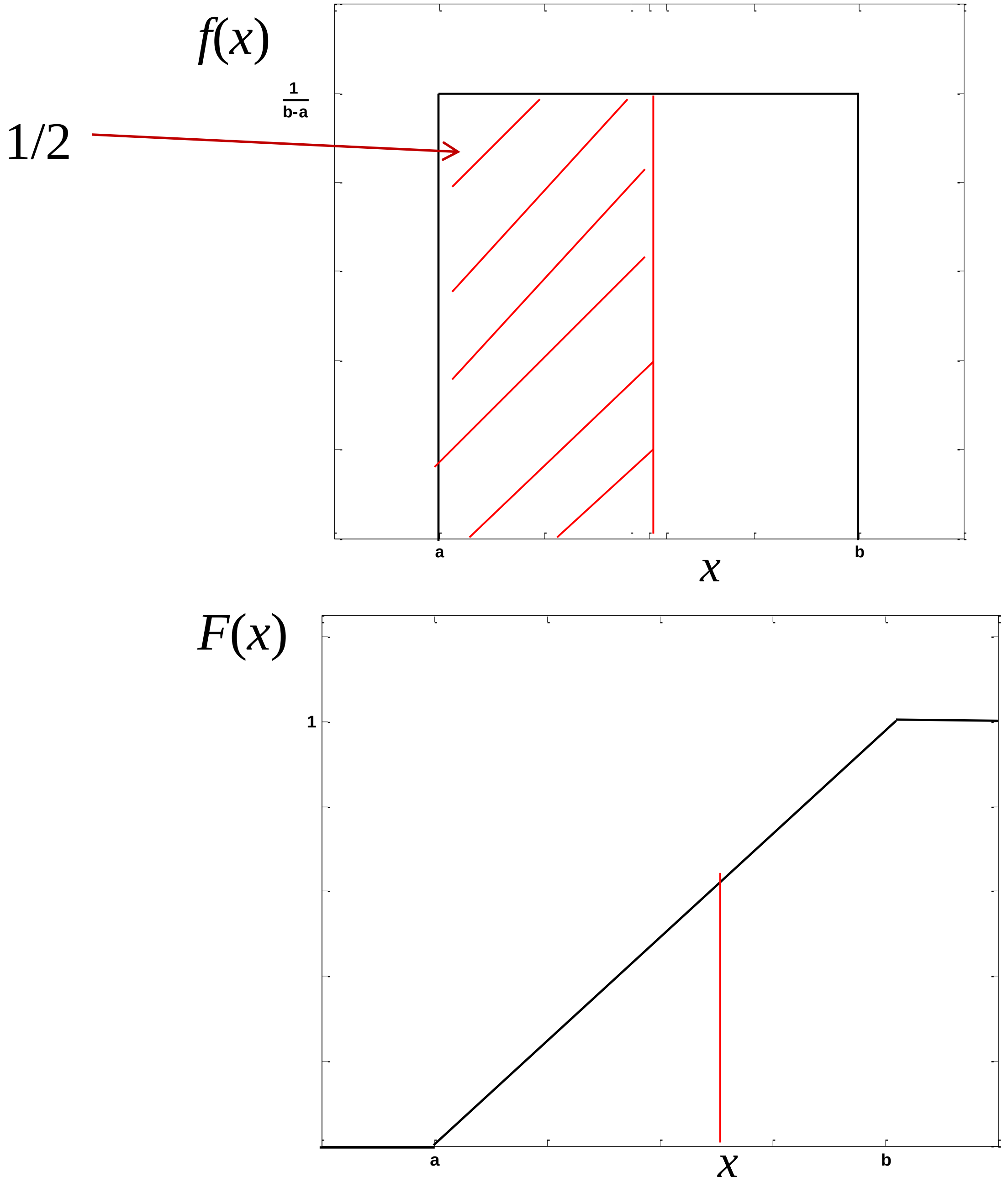
$$\tilde{x} = \frac{b+a}{2}$$

mode

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

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

flat distribution



Uniform and Beta PDF

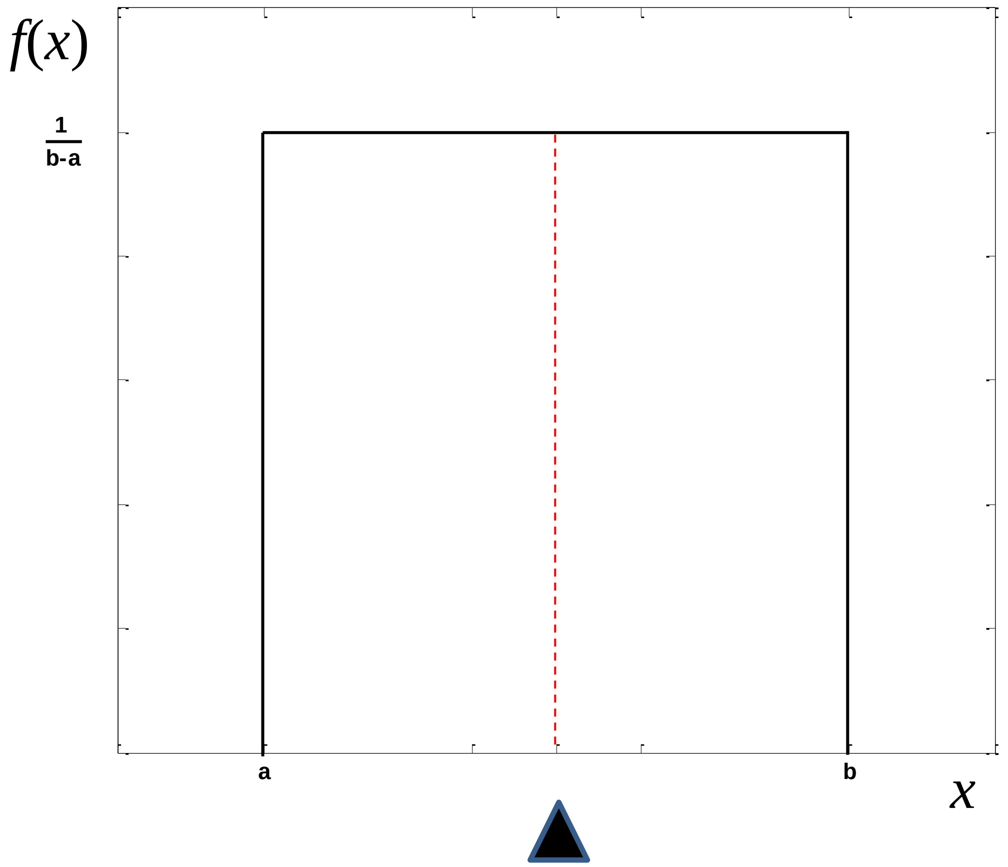
that

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

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

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

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



Uniform and Beta PDF

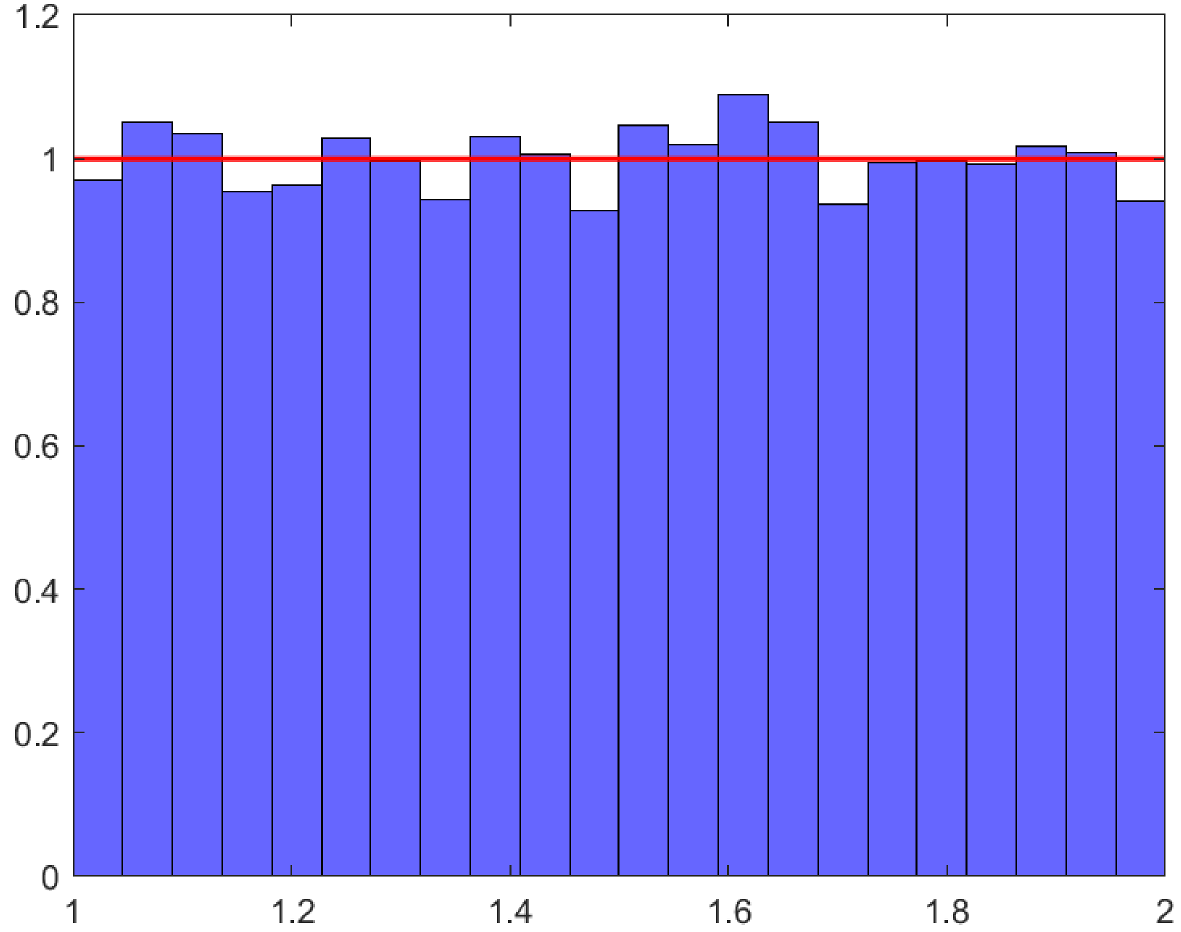
```
rng('default'), a=1;,b=2;,num=10^4;
x=a+(b-a)*rand(num,1);
figure;
histogram(x,'BinLimits',[1,2],'normalization','pdf','FaceColor','blue')
hold on
line([1,2],[1,1],'Color','red','LineWidth',1.5)
xlim([1,2])
```

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

	True	Simulated
μ	1.5	1.4996
σ^2	0.0833	0.0829

Can also find and plot ECDF.

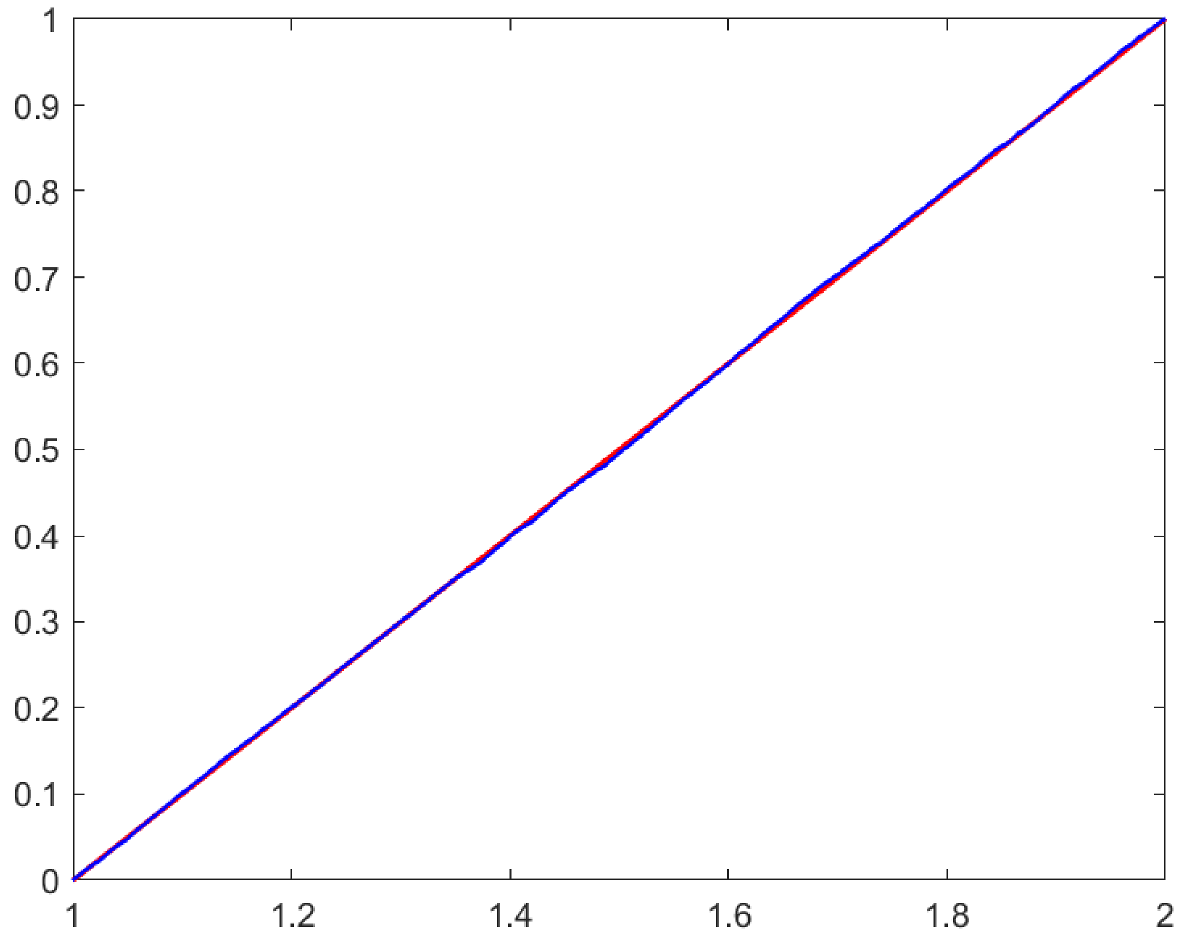
```
themean=(b-a)/2
thevar=(b-a)^2/12
[mean(x),var(x)]
```



Uniform and Beta PDF

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

```
y =cdf('unif',(a:.01:b),a,b);
figure;
plot((1:.01:2),y,'Color','red','LineWidth',1.5)
[F,xx]=ecdf(x);
hold on
stairs(xx,F,'Color','blue','LineWidth',1.5)
```



Uniform and Beta PDF

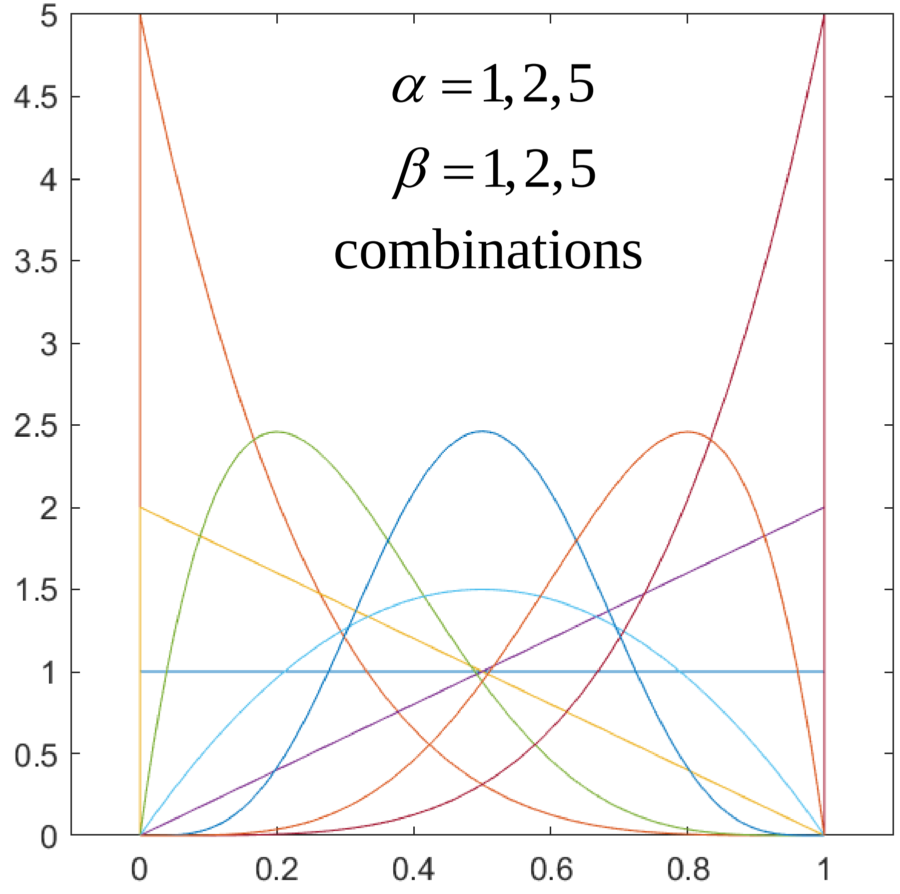
A random variable x has a continuous Beta distribution, x~Beta(α,β) if

$$f(x|\alpha,\beta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1},$$

where, $\alpha,\beta>0$, $x\in[a,b]$.

And can take on many shapes.

$\Gamma(a)=(a-1)!$ for integer a



Uniform and Beta PDF

There is no closed form analytic formula for the CDF $F(x|\alpha,\beta)$.

It can be shown that the mean is

$$\begin{aligned}
\mu &= \int_x x f(x|\theta)dx \\
&= \int_{x=0}^{1} x \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}dx \\
&= \frac{\alpha}{\alpha+\beta}
\end{aligned}$$

Uniform and Beta PDF

It can be shown that the variance is

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

$$= \int_{x=0}^{1} (x-\mu)^2 \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}dx$$

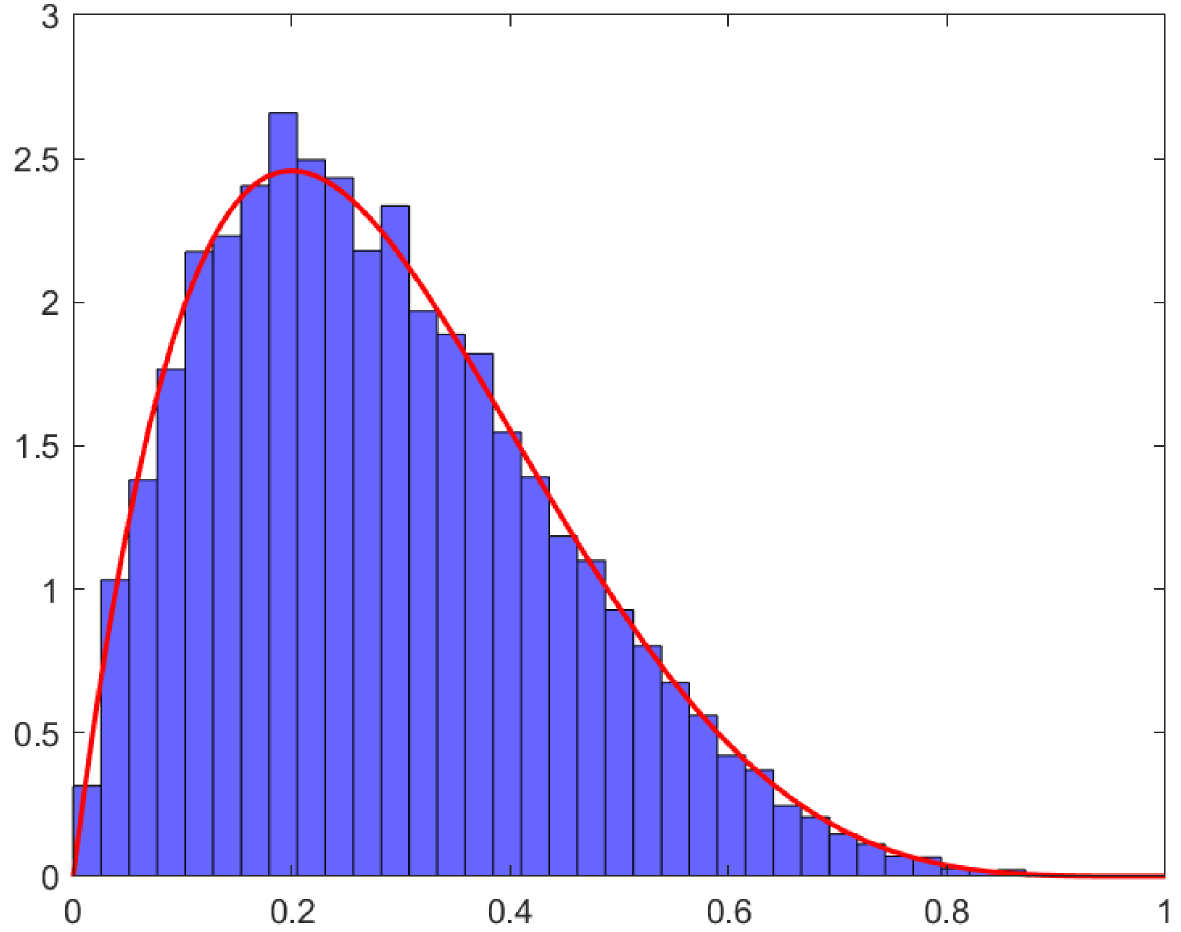
$$= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Uniform and Beta PDF

```
rng('default'), a=2; b=5; num=10^4;
x=betarnd(a,b,[num,1]);
figure;
histogram(x,'BinLimits',[0,1],'normalization','pdf','FaceColor','blue')
hold on
fx = @(x) factorial(a+b-1)/factorial(a-1)/factorial(b-1)...
.*x.^(a-1).*(1-x).^(b-1);
fplot(fx,[0,1],'r','LineWidth',1.5)
```

$$f(x\,|\,\alpha,\beta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$$

	True	Simulated
μ	0.2857	0.2858
σ^2	0.0255	0.0251

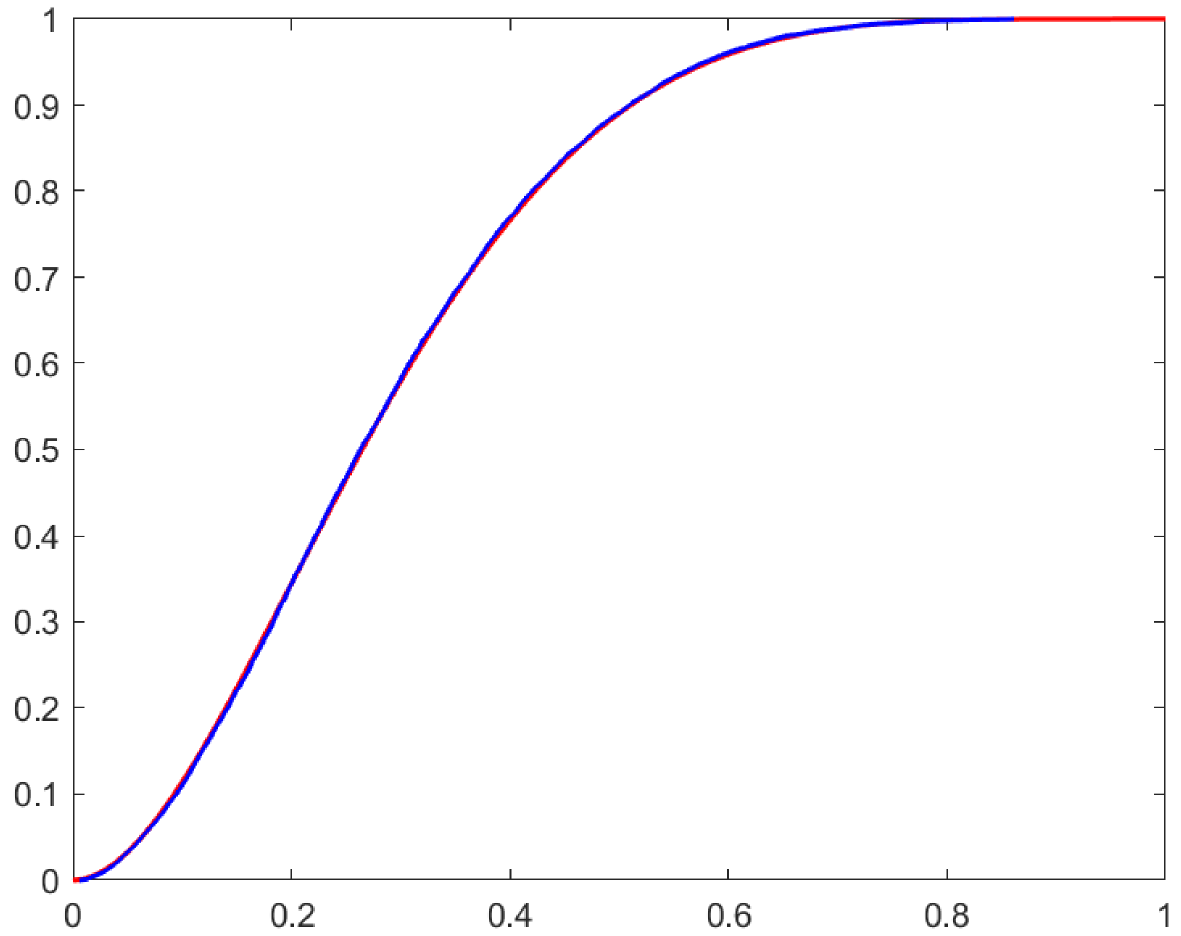


Can also find and plot ECDF.

```
themean=a/(a+b)
thevar=(a*b)/((a+b)^2*(a+b+1))
[mean(x),var(x)]
```

Uniform and Beta PDF

```
y =cdf('beta',(0:.01:1),a,b);
figure;
plot((0:.01:1),y,'Color','red','LineWidth',1.5)
[F,xx]=ecdf(x);
hold on
stairs(xx,F,'Color','blue','LineWidth',1.5)
```

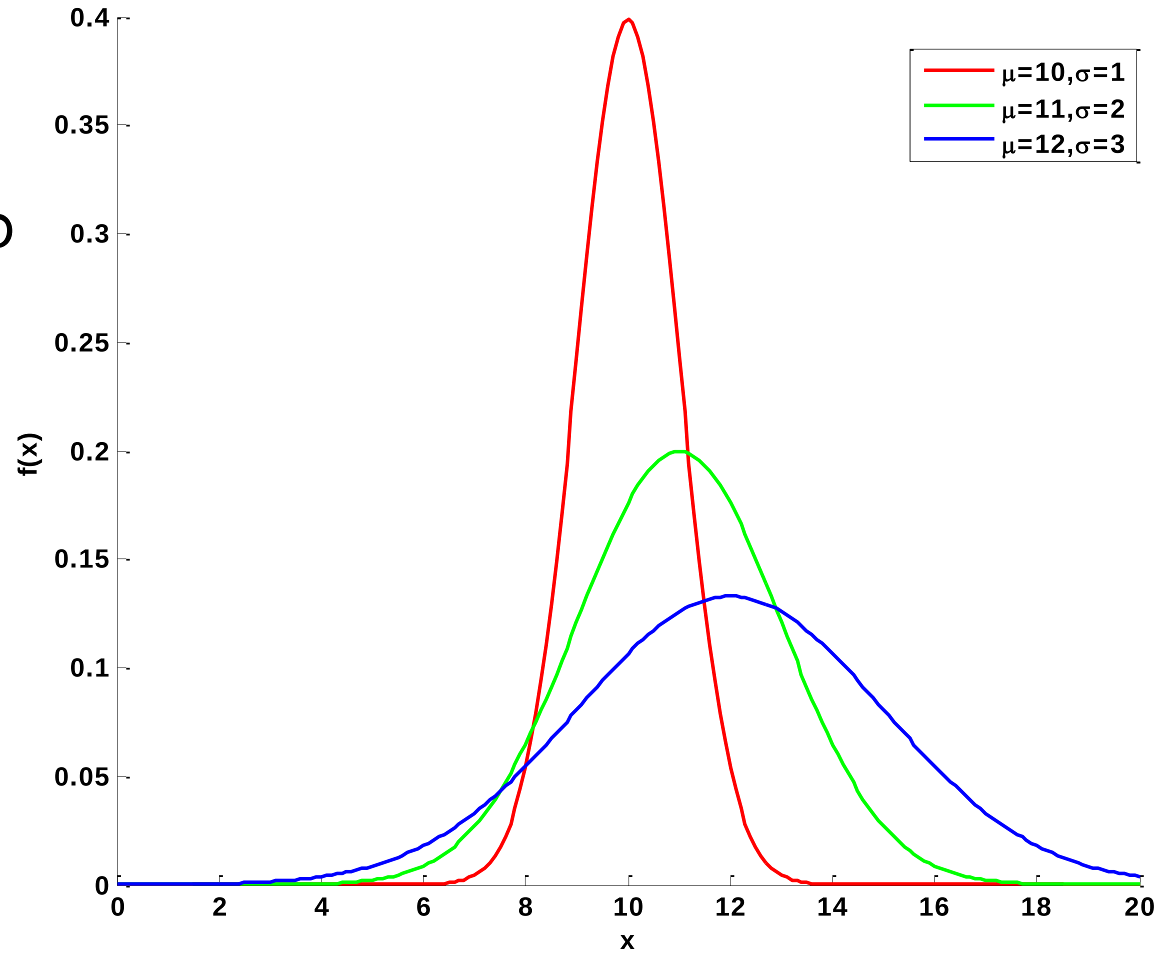


Normal PDF

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

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

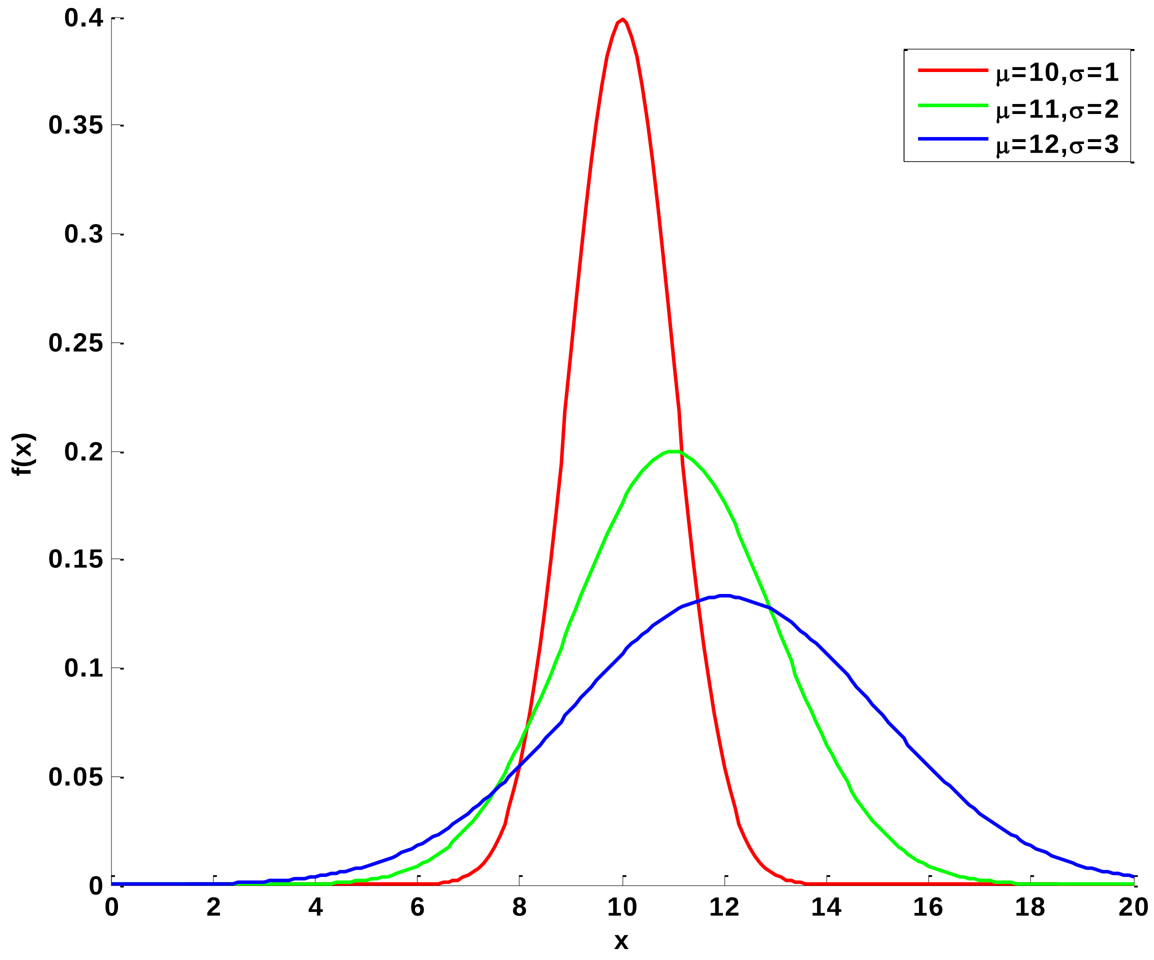
where, $-\infty<\mu<+\infty$ and $\sigma>0$.



Normal PDF

```
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])
```

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



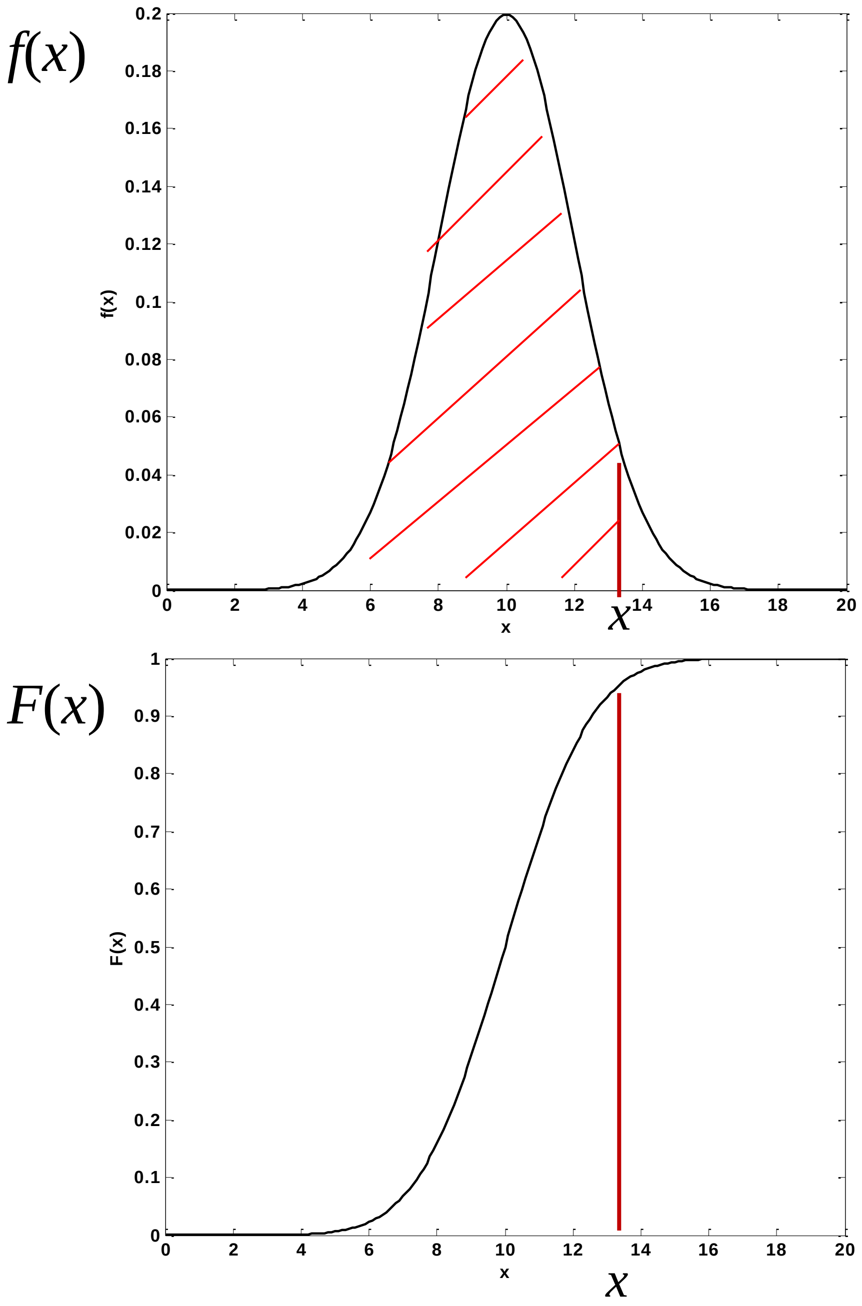
Normal PDF

The CDF of the continuous normal distribution is

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

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

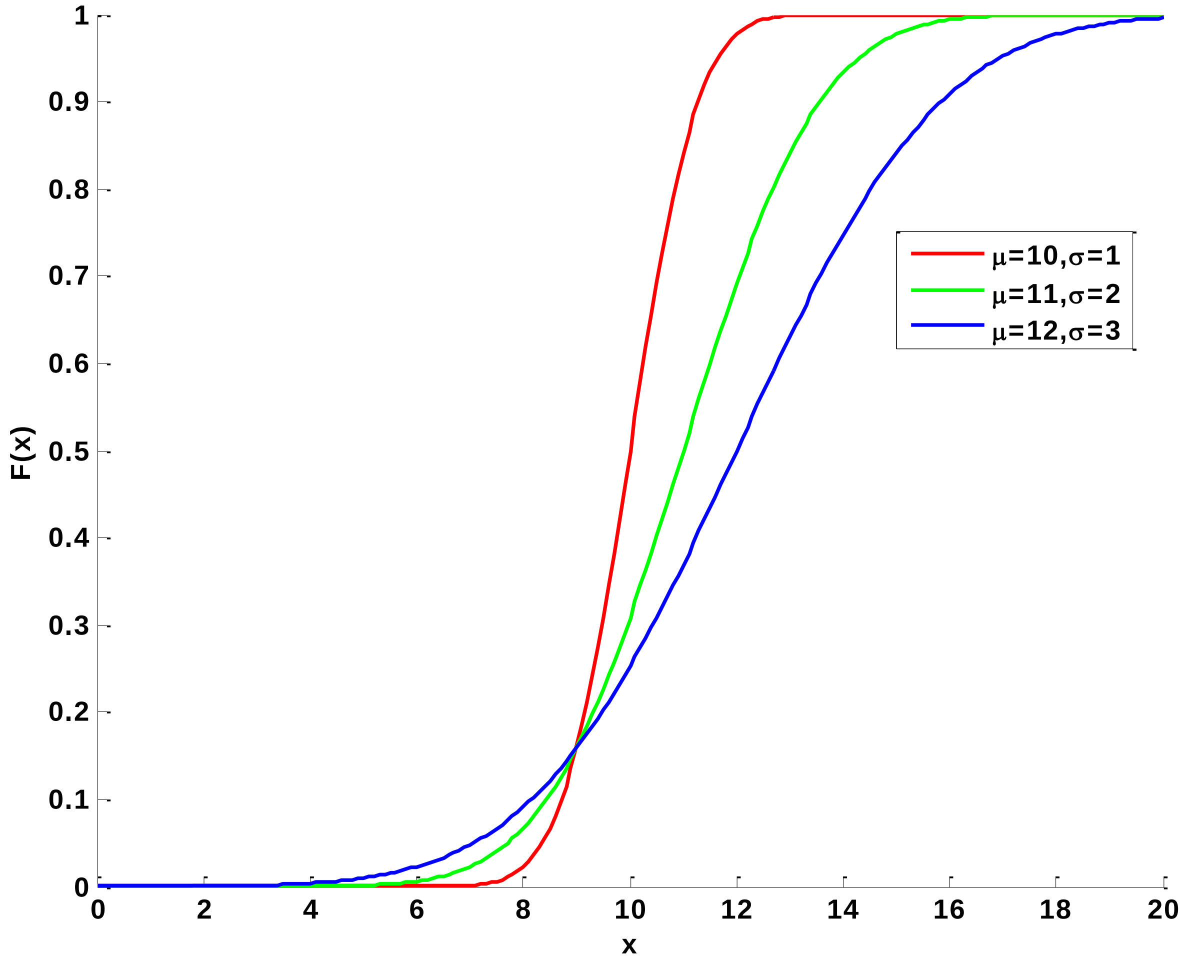
No closed form analytic solution.



Normal PDF

```
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])
```

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

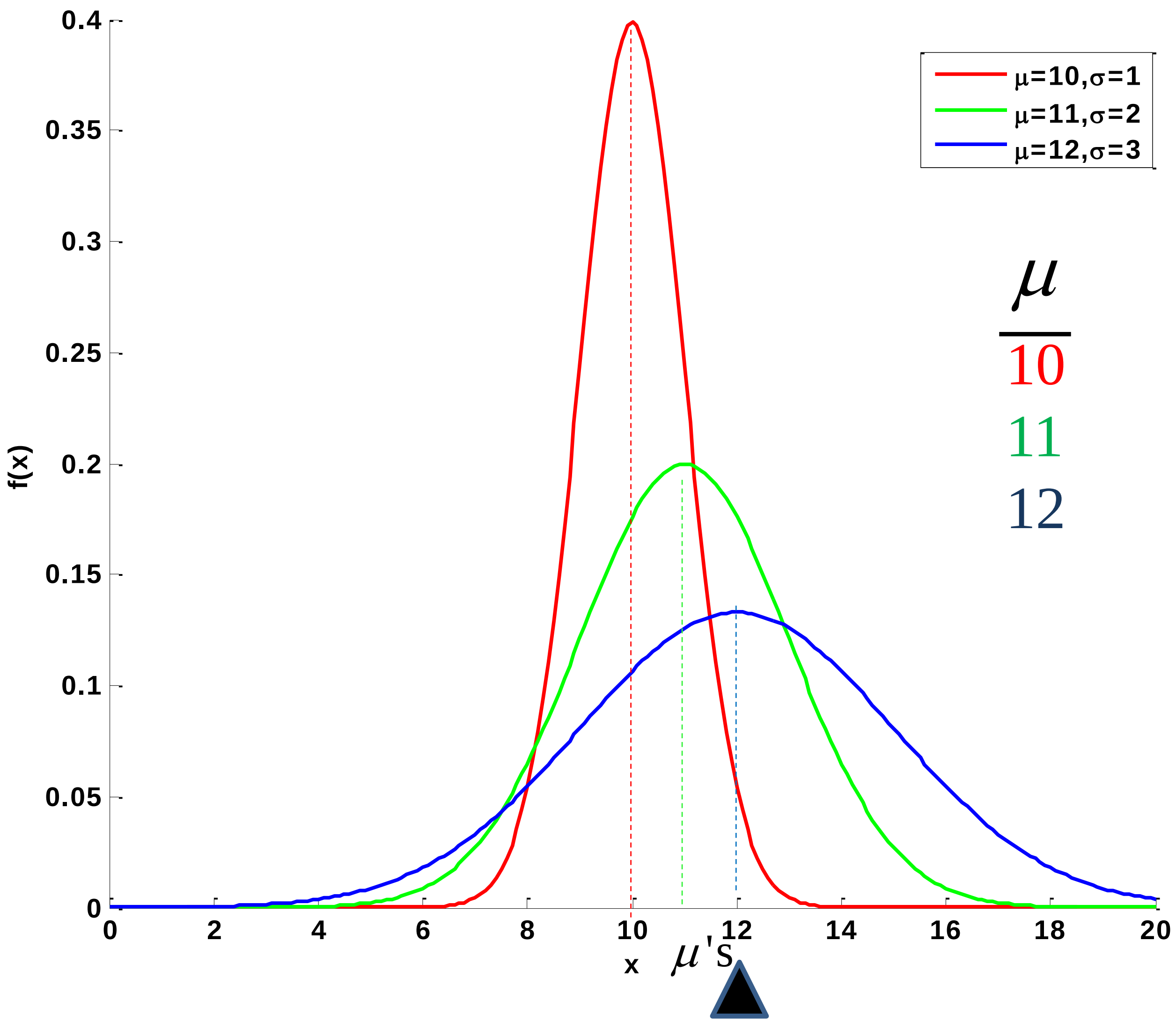


Normal PDF

It can be shown that

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

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



Normal PDF

It can be shown that

median

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

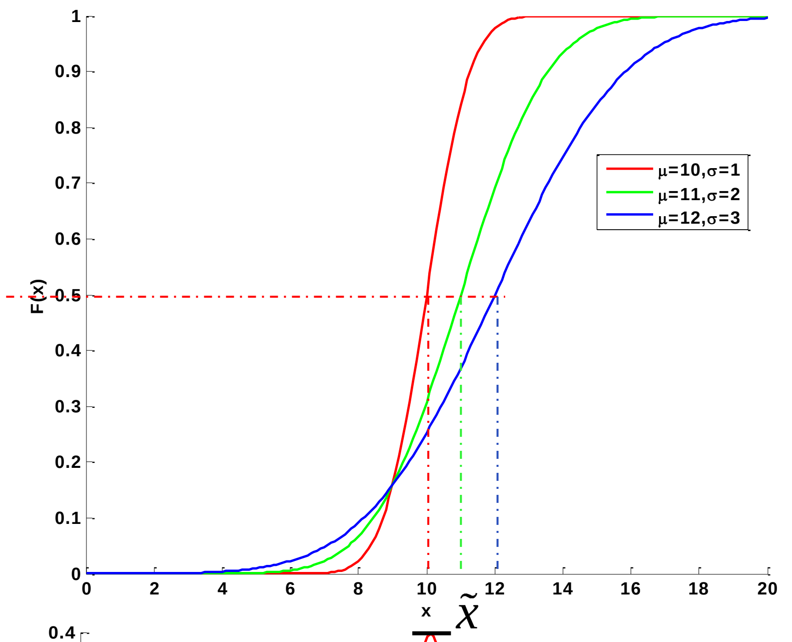
$$\tilde{x} = \mu$$

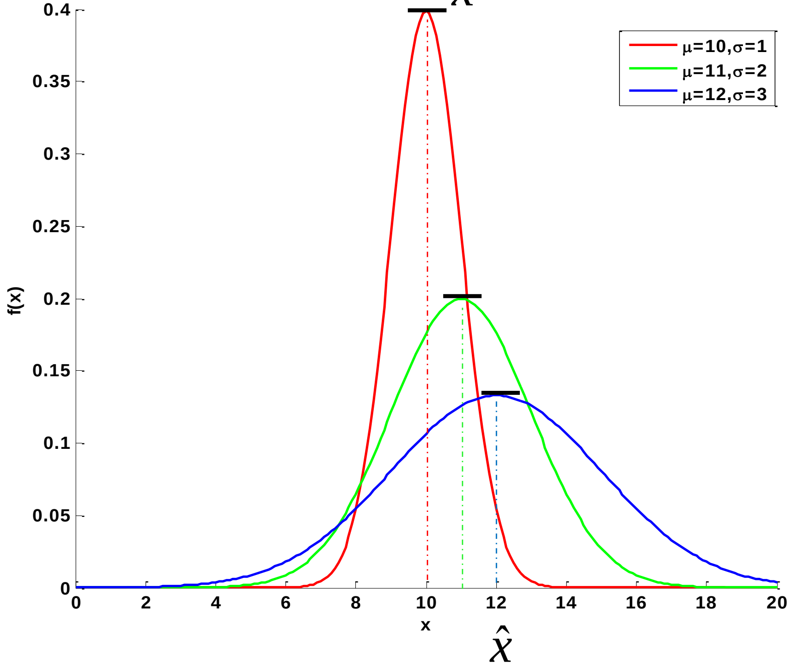
mode

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

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

1/2





$\tilde{x}$
10
11
12

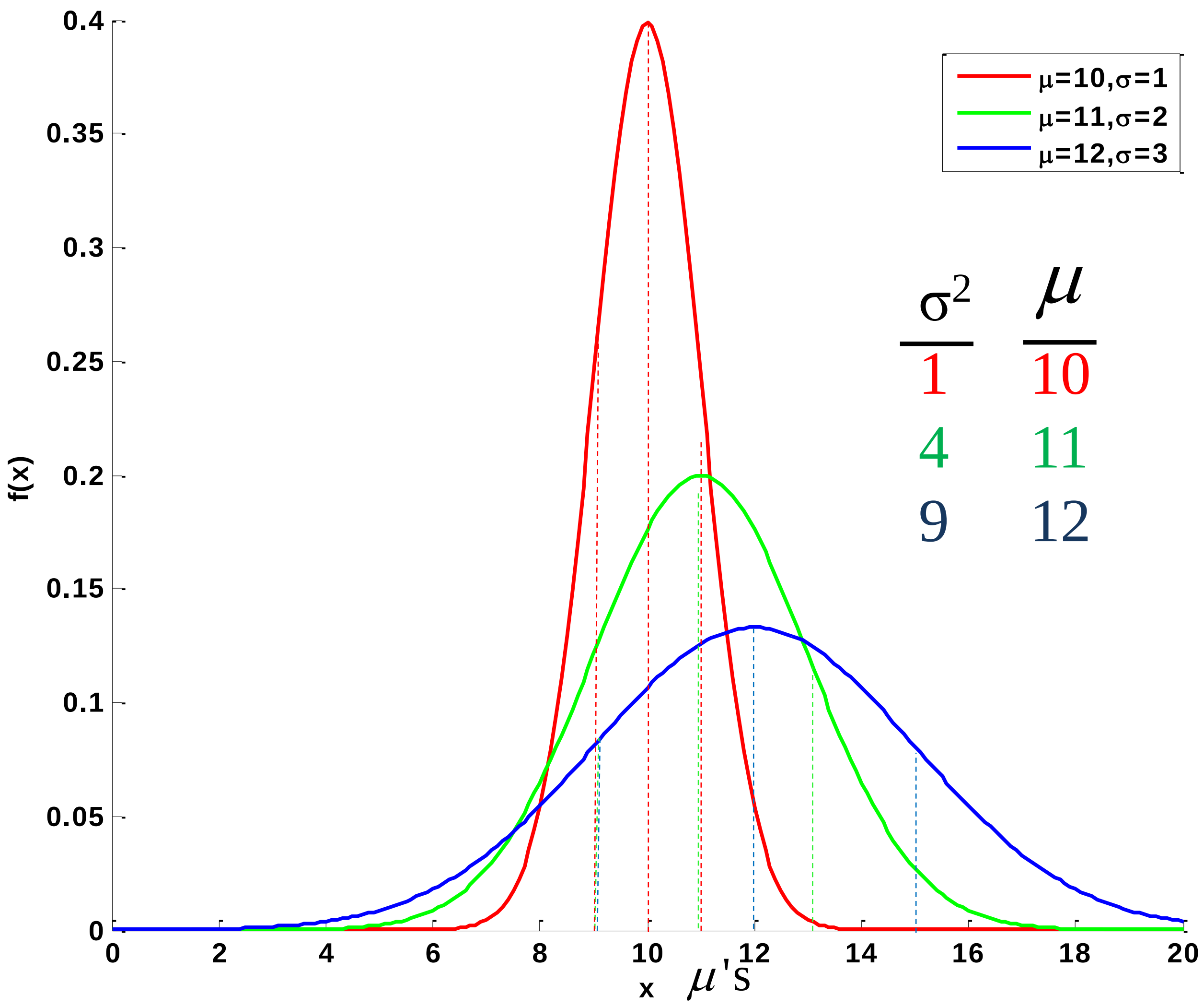
$\hat{x}$
10
11
12

Normal PDF

that

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

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



Normal PDF

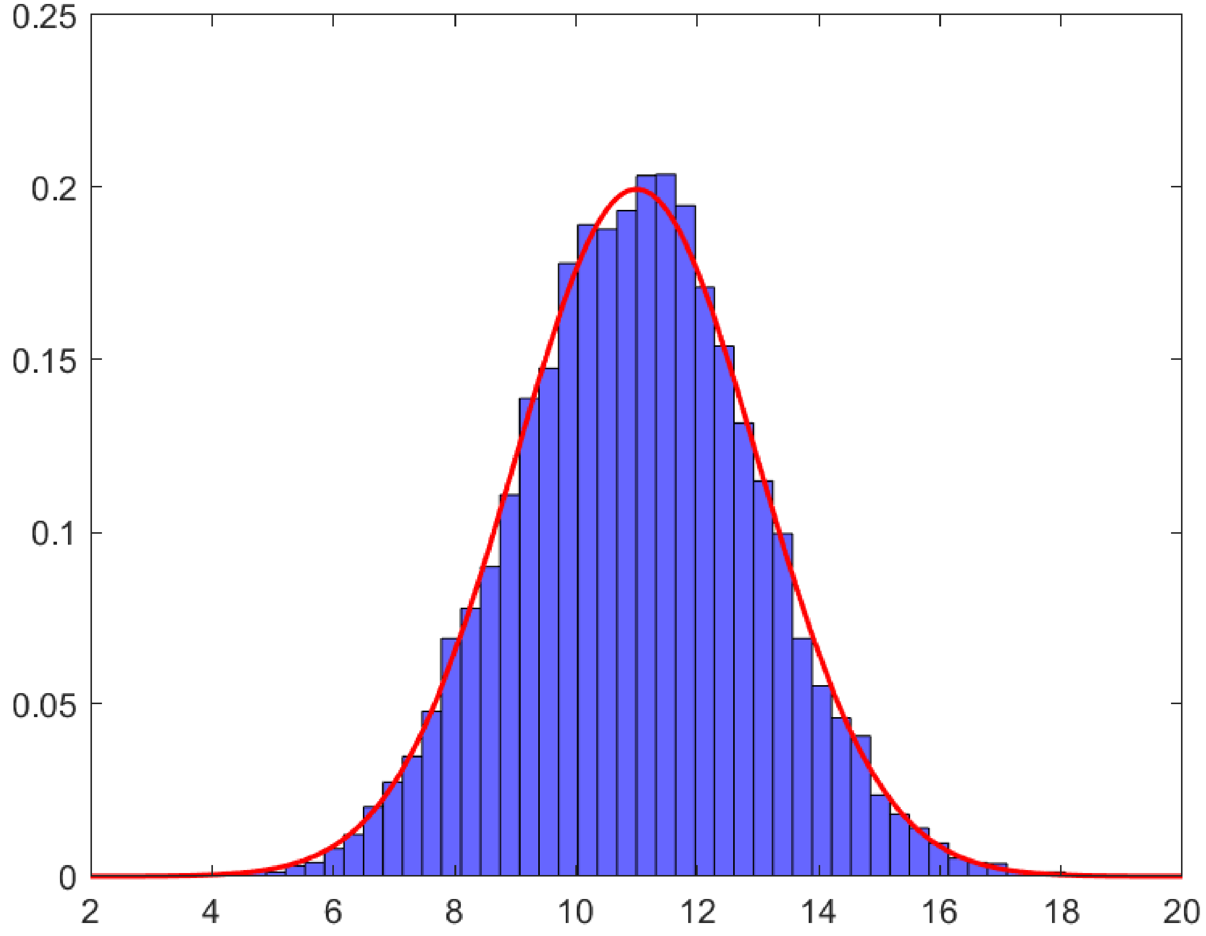
```
rng('default’), mu=11;,sigma=2;,num=10^4;
x=sigma*randn(num,1)+mu;
figure;
histogram(x,'BinLimits',[2,20],'normalization','pdf','FaceColor','blue')
hold on
fx = @(z) 1/sqrt(2*pi*sigma^2).*exp(-(z-mu).^2/(2*sigma^2));
fplot(fx,[0,22],'r','LineWidth',1.5)
xlim([2,20])
```

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

	True	Simulated
μ	11	11.0033
σ^2	4	3.9321

Can also find and plot ECDF.

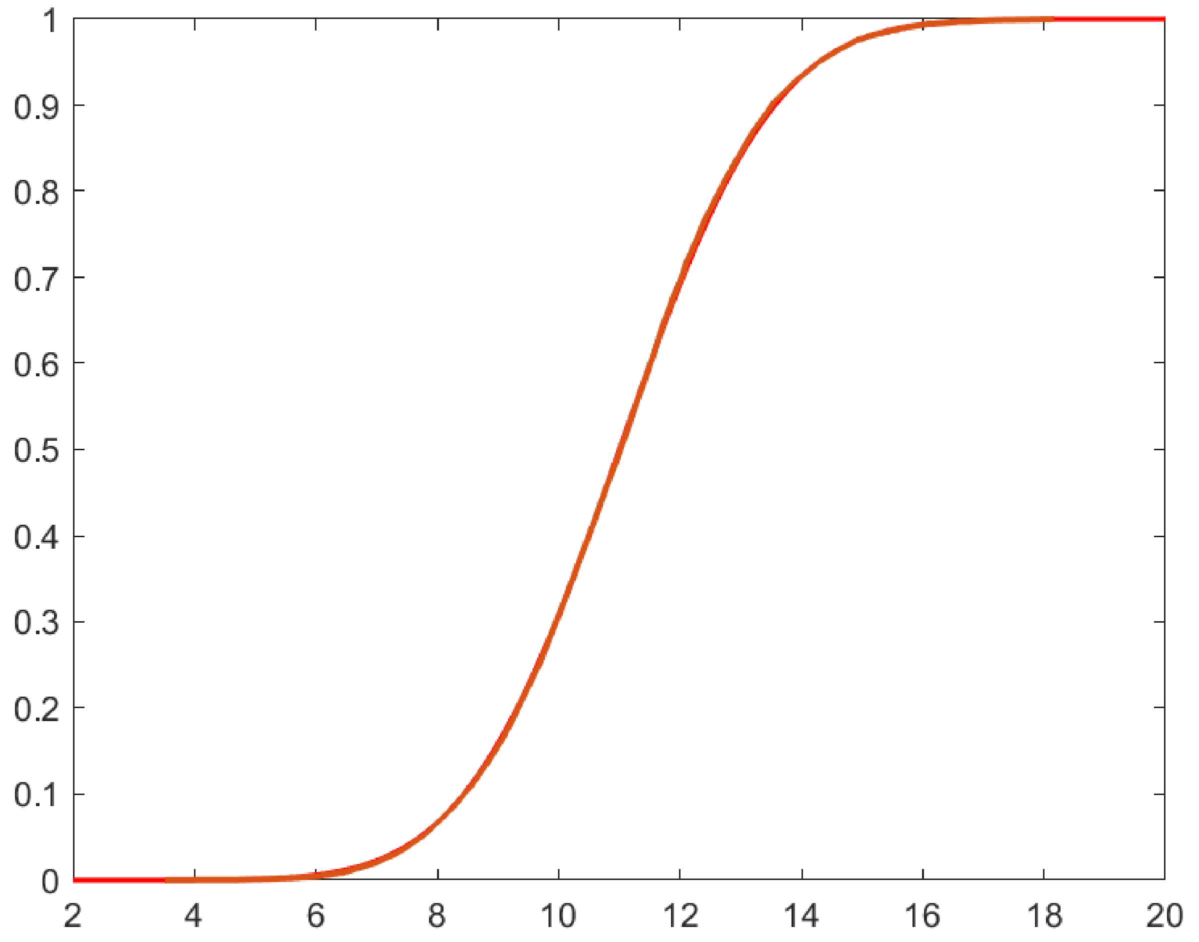
```
themean=mu
thevar=sigma^2
[mean(x),var(x)]
```



Normal PDF

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

```
y =cdf('Normal',(2:.01:20),mu,sigma);
figure;
plot((2:.01:20),y,'Color','red','LineWidth',1.5)
[F,xx]=ecdf(x);
hold on
stairs(xx,F,'Color','blue','LineWidth',1.5)
```



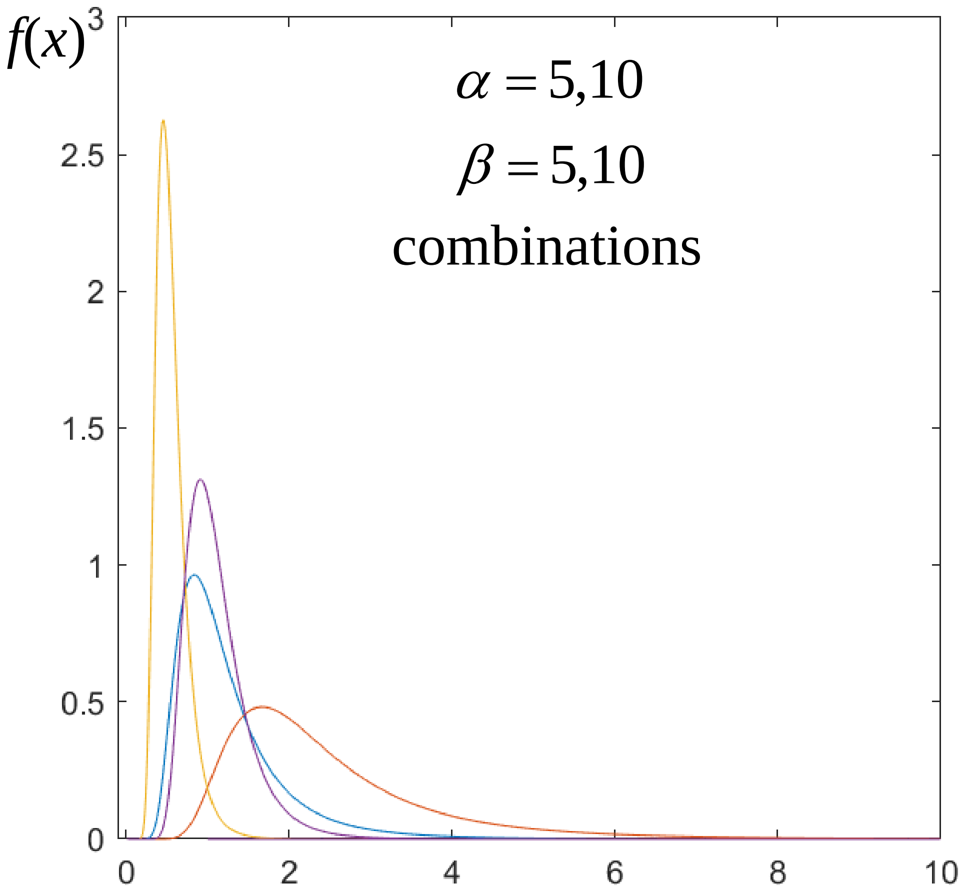
Inverse Gamma PDF

A random variable x has a continuous Inverse Gamma distribution, $x \sim \mathrm{Inv}\Gamma(\alpha,\beta)$ if

$$f(x|\alpha,\beta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{-\alpha-1}e^{-\beta/x} \quad ,$$

where, $\alpha,\beta>0$, $0<x<\infty$.

And can take on many shapes.



$\Gamma(a)=(a-1)!$ for integer a

Inverse Gamma PDF

There is no closed form analytic formula for the CDF $F(x|\ \alpha,\beta)$.

It can be shown that the mean is

$$\begin{aligned}\mu &= \int_x xf(x\,|\,\theta)dx \\ &= \int_{x=0}^{\infty} x\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{-\alpha-1}\mathrm{e}^{-\beta/x}\,dx \\ &= \frac{\beta}{\alpha-1}\end{aligned}$$

Inverse Gamma PDF

It can be shown that the variance is

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

$$= \int_{x=0}^{\infty} (x-\mu)^2 \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} e^{-\beta/x} dx$$

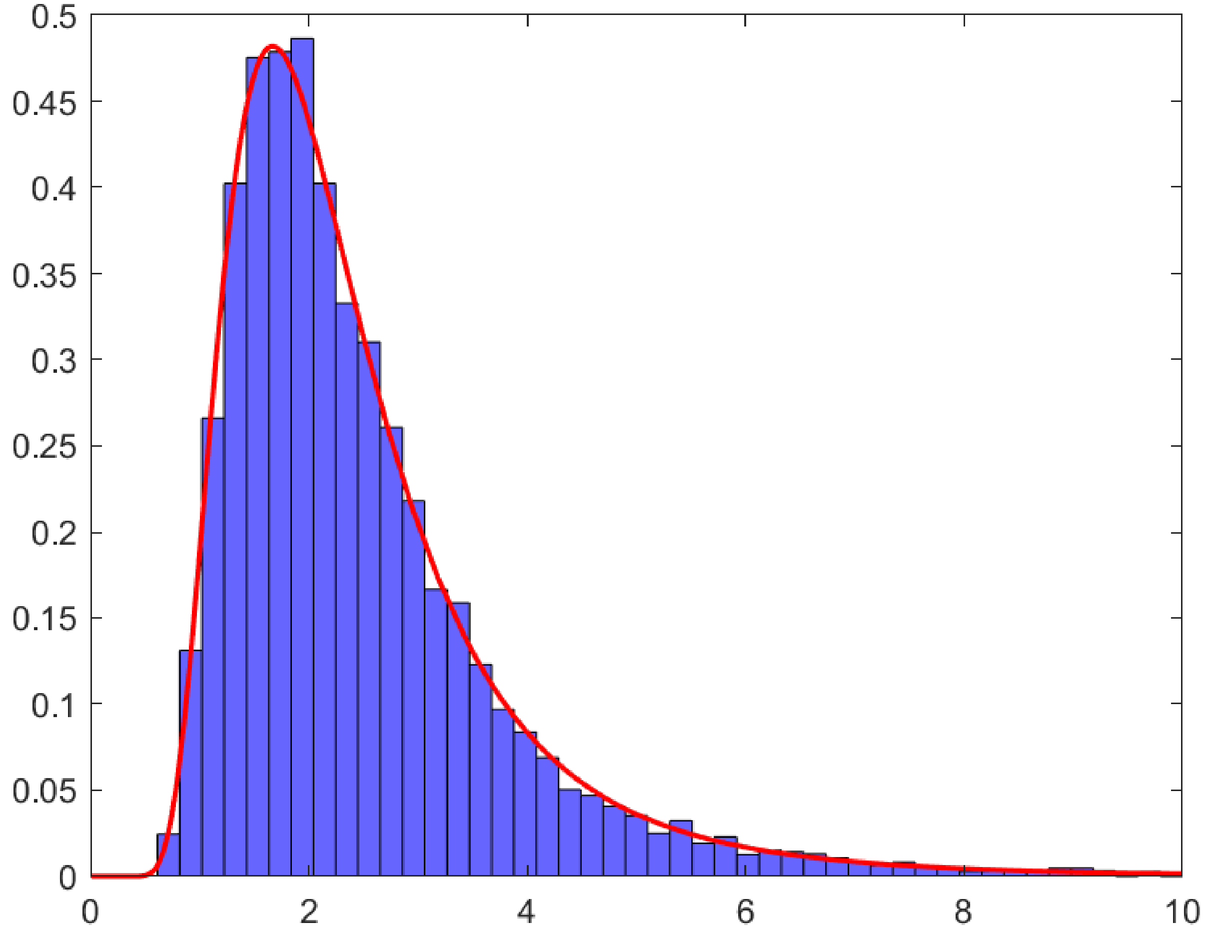
$$= \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$$

Inverse Gamma PDF

```
rng('default'), a=5; b=10; c=1/b; num=10^4;
x=1./gamrnd(a,c,[num,1]);
figure;
histogram(x,'BinLimits',[0,10],'normalization','pdf','FaceColor','blue')
hold on
fx = @(z) b.^a/factorial(a-1).*z.^(-a-1).*exp(-b./z);
fplot(fx,[0,10],'r','LineWidth',1.5)
xlim([0,10])
```

$$f(x|\alpha,\beta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{-\alpha-1}\mathrm{e}^{-\beta/x}$$

	True	Simulated
μ	2.5000	2.4877
σ^2	2.0833	1.9690

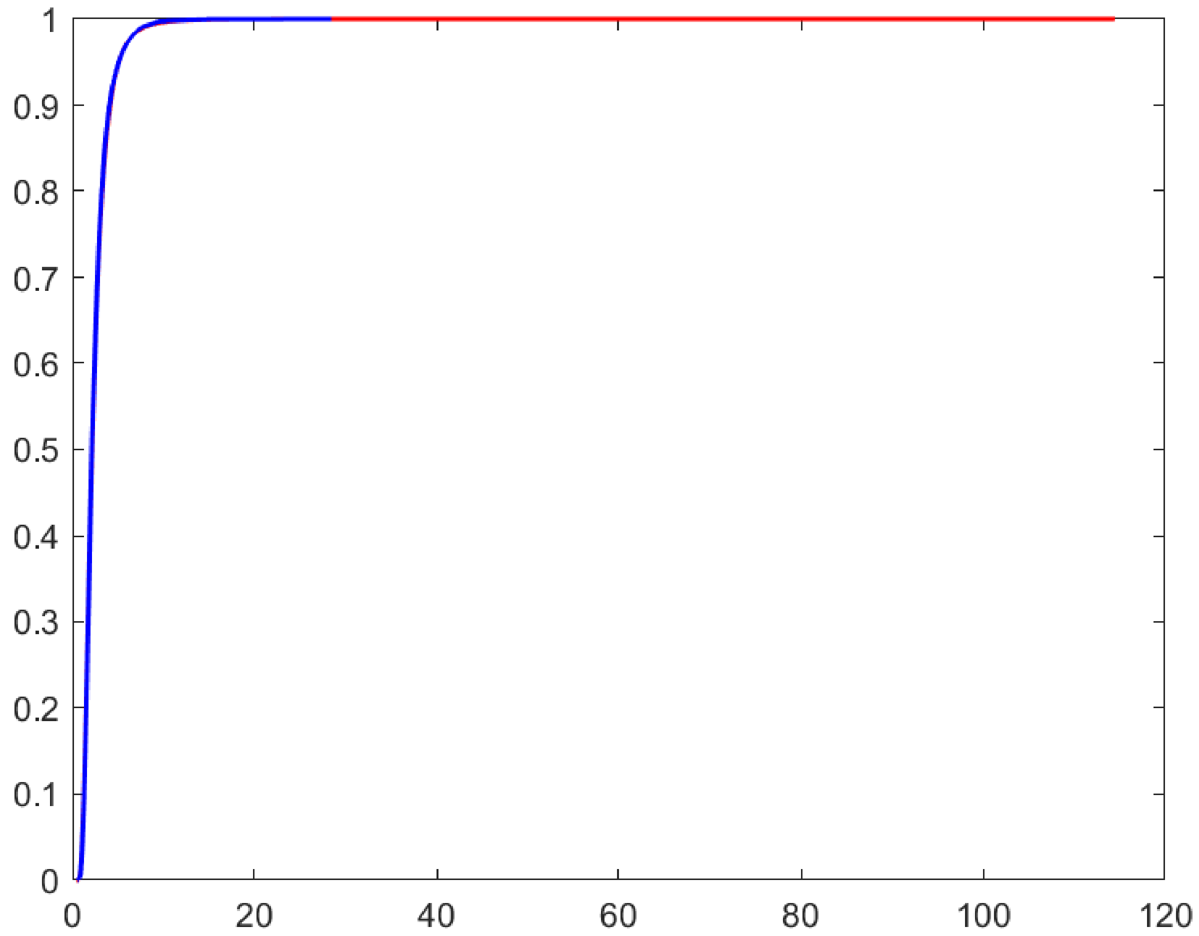


Can also find and plot ECDF.

```
themean=b/(a-1)
thevar=(b^2)/((a-1)^2*(a-2))
[mean(x),var(x)]
```

Inverse Gamma PDF

```
y=1./gamrnd(a,c,[5*10^6,1]);
[Fy,yy]=ecdf(y);
figure;
stairs(yy,Fy,'Color','red','LineWidth',1.5)
[F,xx]=ecdf(x);
hold on
stairs(xx,F,'Color','blue','LineWidth',1.5)
```



Discussion

Questions?

Homework 3

$$f(x|\alpha,\beta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$$

1. Use Matlab to generate n=1000 observations from the Beta PDF with α=3 and β=7. Change α,β to observe shapes.

a. Calculate the sample mean.

b. Calculate the sample variance.

c. Calculate the median.

d. Make a histogram.

e. Make an empirical CDF and calculate the 95$^{\text{th}}$ percentile

Homework 3

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

2. Use Matlab to generate n=1000 observations from the normal PDF with μ=100 and σ^2=5.

a. Calculate the sample mean.

b. Calculate the sample variance.

c. Calculate the median.

d. Make a histogram.

e. Make an empirical CDF and calculate the 95^{th} percentile.

Homework 3

$$f(x|\alpha,\beta)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}x^{\alpha-1}\mathrm{e}^{-\beta x}$$

3. Assume y has a gamma distribution with parameters with $\alpha=3$, $\beta=7$, and $c=1/\beta$. Change α,β to observe shapes.

a. Generate $n=1000$ random y observations.
Calculate the sample mean and variance.
Make a histogram. Compare to the Gamma PDF.

b*. Using pencil and paper transform to x as x=1/y.

c. Take $x=1/y$ for each x random observation in part a.

d. Repeat a. but now for y's and compare to inverse gamma PDF.

*For students in MSSC 5790