ELEG-667

Statistical Signal Processing

Prof. Barner

316 Evans Hall
831-6937
barner@udel.edu

Goal: Given a discrete time sequence $\{X(n)\}$, how we develop

- Statistical and spectral representations
- Filtering, prediction, and system identification algorithms
- Optimization methods
  - Statistical
  - Adaptive

Many methods assume that $\{x(n)\}$ is deterministic. Real world signals are usually statistical in nature.

Thus,

$$\ldots\ x(-1),\ x(0),\ x(1),\ \ldots$$

can be interpreted as a sequence of random variables.

- We begin by analyzing each observation $x(n)$ as a *R.V.*
- Then, to capture dependencies, we consider random vectors

$$\ldots\ \underbrace{x(n),\ x(n+1),\ \ldots,\ x(n+N-1)}_{\underset{\sim}{x}(n)},\ x(n+N),\ \ldots$$

Random Variables

For a space *S*, the subsets, or events of *S* have associated probabilities. To every event δ, we assign a number $x(\delta)$, which is called a *R.V.*

The distribution function of x is

$$\Pr\{x \le x_0\} = F_x(x_0) \qquad -\infty < x_0 < \infty$$

Properties:

1) $F(+\infty) = 1, \qquad F(-\infty) = 0$

2) $F(x)$ is continuous from the right

$$F(x^+) = F(x)$$

3) $\Pr\{x_1 < x \le x_2\} = F(x_2) - F(x_1)$

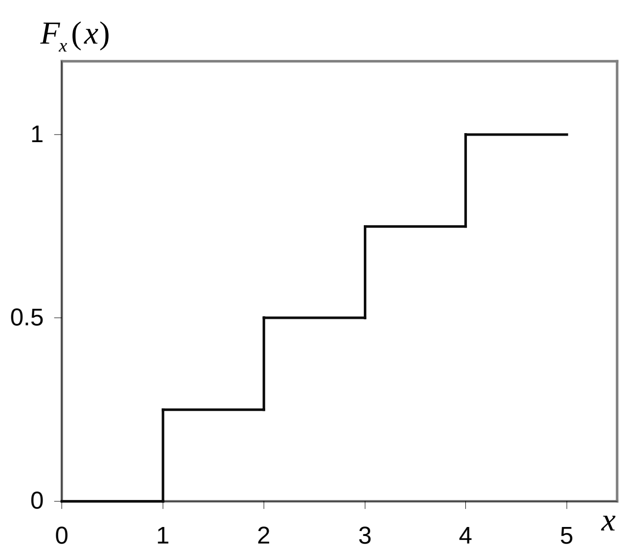
Example: fair toss of two coins

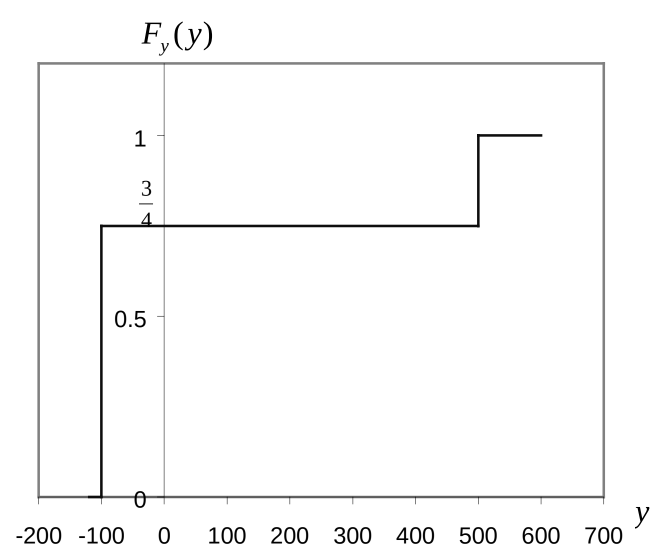
Events(δ)	*Prob.*	*X(δ)*	*Y(δ)*
HH	¼	1	-100
HT	¼	2	-100
TH	¼	3	-100
TT	¼	4	500

This yields different distribution functions

$$F_x(2) = \Pr\{HH, HT\} = 1/2$$

$$F_y(2) = \Pr\{HH, HT, TH\} = 3/4$$





The probability density function is defined as,

$$f(x) = \frac{dF(x)}{dx}$$

or $$F(x) = \int_{-\infty}^{x} f(x)dx$$

Thus $$F(\infty) = 1 \Rightarrow \int_{-\infty}^{\infty} f(x)dx = 1$$

Types of distributions:

- Continuous: $\Pr\{x = x_0\} = 0 \qquad \forall x_0$
- Discrete: $F(x_i) - F(x_i^{-}) = \Pr\{x = x_i\} = P_i$

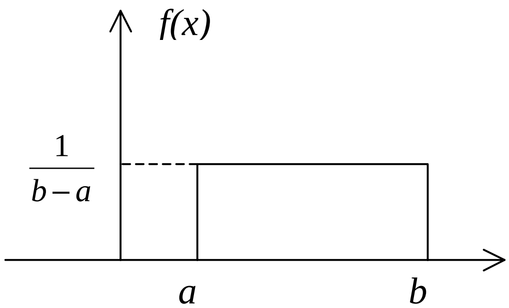
  In which case $f(x) = \sum_i P_i \delta(x - x_i)$

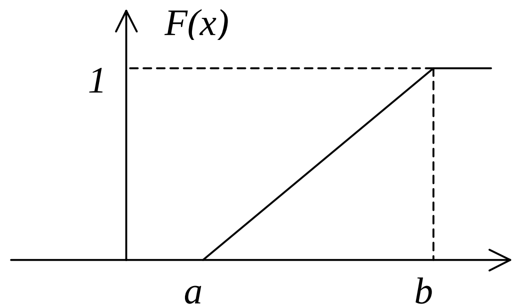
- Mixed: discontinuous but not discrete.

Distribution examples

Uniform: $x \sim U(a,b) \qquad a < b$

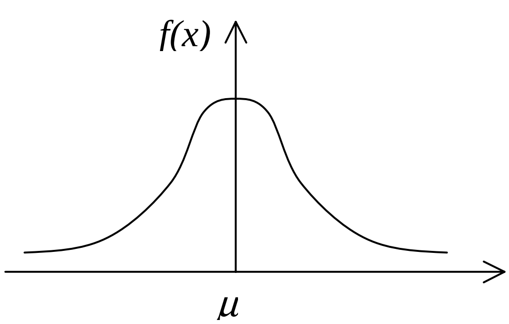
$$f(x) = \begin{cases} \dfrac{1}{b-a} & x \in [a,b] \\ 0 & else \end{cases}$$

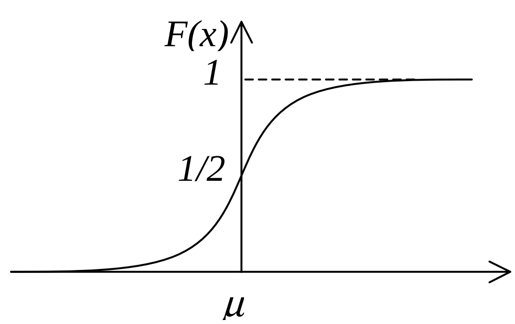




Gaussian: $x \sim N(\mu, \sigma)$

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





Binomial:

Example: Toss a coin n times. What is the probability of getting k heads?

For $p+q=1$

Probability of tail

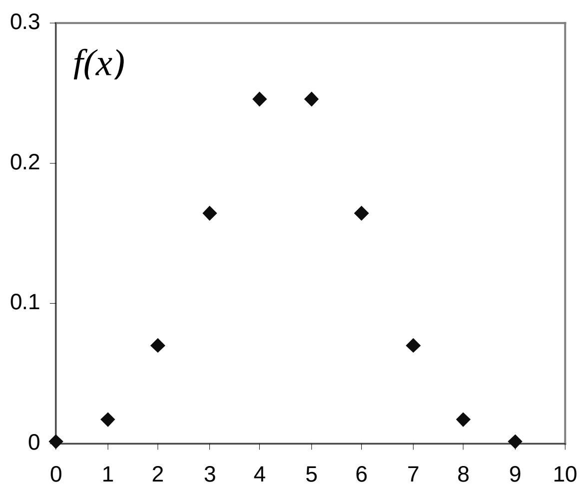
Probability of head

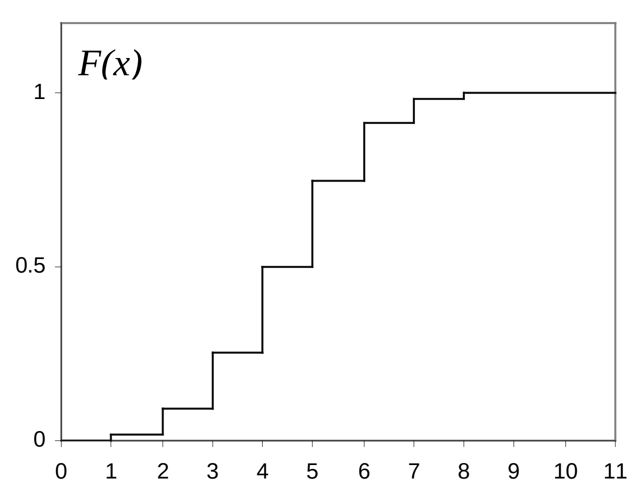
$$\Pr\{x=k\}=\binom{n}{k}p^k q^{n-k}$$

$$\Rightarrow f(x)=\sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k}\delta(x-k)$$

$$\Rightarrow F(x)=\sum_{k=0}^{m}\binom{n}{k}p^k q^{n-k} \qquad m\le x<m+1$$

For $n=9$, $p=q=1/2$,





Conditional Distributions

The conditional distribution of *x* given event “M” has occurred is

$$F_x(x_0 \mid M) = \Pr\{x \le x_0 \mid M\}$$
$$= \frac{\Pr\{x \le x_0, M\}}{\Pr\{M\}}$$

Example: Suppose $M = \{x \le a\}$

Then

$$F_x(x_0 \mid M) = \frac{\Pr\{x \le x_0, x \le a\}}{\Pr\{x \le a\}}$$

If $x_0 \ge a$, what happens?

In this case

$$\Pr\{x \le x_0, x \le a\} = \Pr\{x \le a\}$$

and

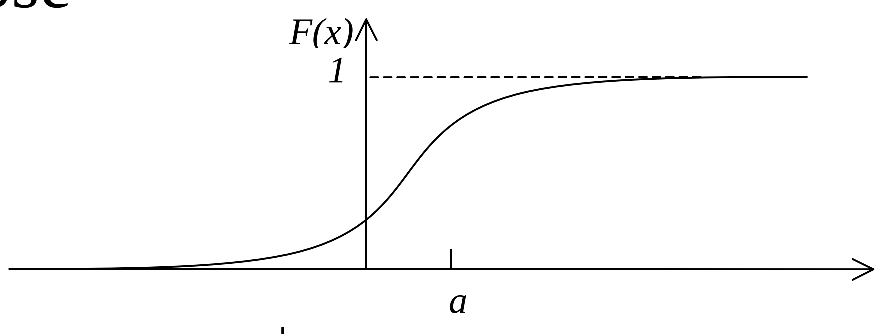
$$F_x(x_0|M) = \frac{\Pr\{x \le a\}}{\Pr\{x \le a\}} = 1$$

If $x_0 \le a$, then

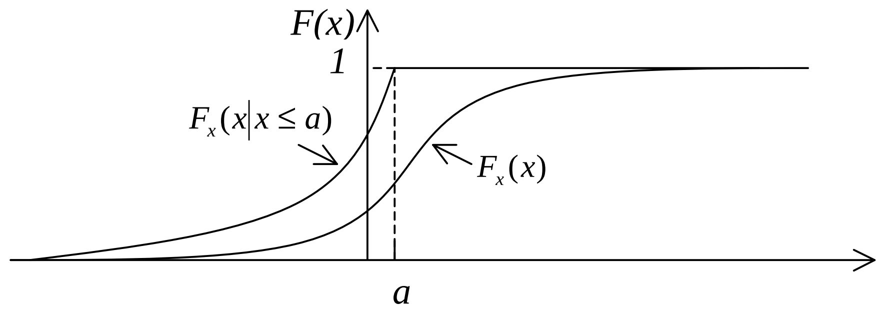
$$F_x(x_0|M) = \frac{\Pr\{x \le x_0\}}{\Pr\{x \le a\}}$$

$$= \frac{F_x(x_0)}{F_x(a)}$$

Suppose



What does $F_x(x|M)$ look like?



As before

$$F(\infty|M) = 1 \qquad F(-\infty|M) = 0$$

$$\Pr\{x_0 \le x \le x_1|M\} = F(x_1|M) - F(x_0|M)$$

and

$$f(x|M) = \frac{\partial F(x|M)}{\partial x}$$

$$F(x_0|M) = \int_{-\infty}^{x_0} f(x|M)dx$$

Example: Toss a fair coin four times, let x be the number of heads.

Recall

$$\Pr\{x = k\} = \binom{n}{k} p^k q^{n-k}$$

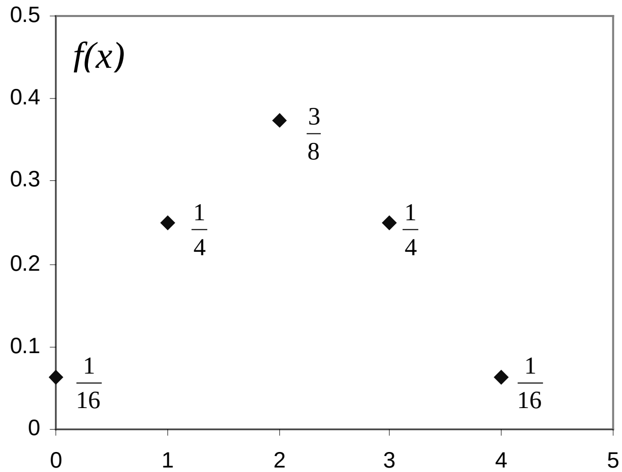
In this case

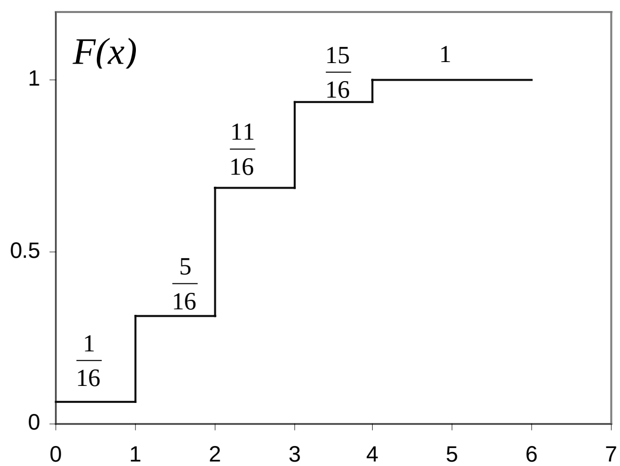
$$\Pr\{x = k\} = \binom{4}{k}\left(\frac{1}{2}\right)^4$$

$$\Pr\{x = 0\} = \Pr\{x = 4\} = 1/16$$

$$\Pr\{x = 1\} = \Pr\{x = 3\} = 1/4$$

$$\Pr\{x = 2\} = 3/8$$





Suppose M = at least one flip produces a head

$$\Pr\{\ M\ \} = 1 - \Pr\{\ \text{No heads}\}$$
$$= 1 - \frac{1}{16}$$
$$= \frac{15}{16}$$

What is $\Pr\{x = k \mid M\}$

$$\Pr\{x = 0 \mid M\} = \frac{\Pr\{x = 0, M\}}{\Pr\{M\}} = 0$$

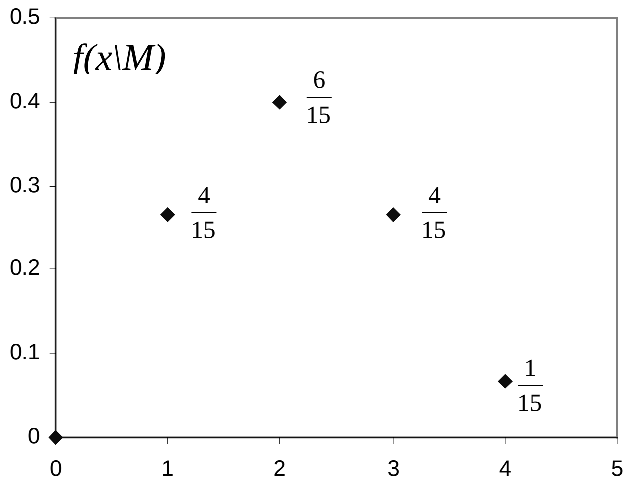
$$\Pr\{x = 1 \mid M\} = \frac{\Pr\{x = 1, M\}}{\Pr\{M\}}$$

$$= \frac{\Pr\{x = 1\}}{\Pr\{M\}} = \frac{1/4}{15/16} = \frac{4}{15}$$

$$\Pr\{x = 2 \mid M\} = \frac{\Pr\{x = 2\}}{\Pr\{M\}} = \frac{3/8}{15/16} = \frac{6}{15}$$

$$\Pr\{x = 3 \mid M\} = \frac{4}{15}$$

$$\Pr\{x = 4 \mid M\} = \frac{1}{15}$$



Total Probability and Bayes' Theorem

Let $M_1, M_2, \ldots, M_n$ forms a partition of S, i.e.

$$\bigcup_i M_i = S \quad \text{and} \quad M_i \bigcap_{i \neq j} M_j = \phi$$

Then

$$F(x) = \sum_i F_x(x \mid M_i) \Pr(M_i)$$

$$f(x) = \sum_i f(x \mid M_i) \Pr(M_i)$$

$$\Pr\{A \mid B\} = \frac{\Pr\{A, B\}}{\Pr\{B\}}$$

$$= \frac{\Pr\{B \mid A\} \Pr(A)}{\Pr\{B\}}$$

From this we get

$$\Pr\{M \mid x \le x_0\} = \frac{F(x_0 \mid M)\Pr(M)}{F(x_0)}$$

and

$$\Pr\{M \mid x = x_0\} = \frac{f(x_0 \mid M)\Pr(M)}{f(x_0)}$$

by integration

$$\Pr\{M\} = \int_{-\infty}^{\infty} \Pr\{M \mid x = x_0\} f(x_0)\,dx_0$$

Bayes' Theorem:

$$f\{x_0 \mid M\} = \frac{\Pr(M \mid x = x_0) f(x_0)}{\Pr(M)}$$

$$= \frac{\Pr(M \mid x = x_0) f(x_0)}{\int_{-\infty}^{\infty} \Pr\{M \mid x = x_0\} f(x_0)\,dx_0}$$

Functions of a R.V.

Let x and $g(x)$ be RVs such that

$$y = g(x)$$

Then

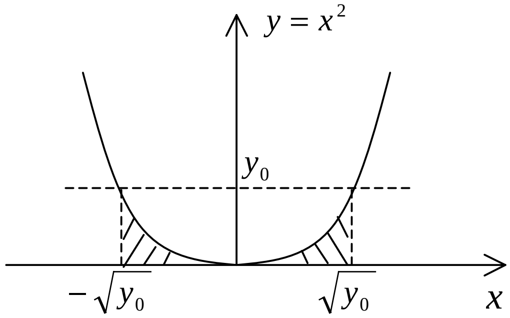
$$\begin{aligned} F_y(y_0) &= \Pr\{ y \le y_0 \} \\ &= \Pr\{ g(x) \le y_0 \} \\ &= \Pr\{ x \in R_{y_0} \} \end{aligned}$$

Where

$$R_{y_0} = \{ x : g(x) \le y_0 \}$$

If $y = g(x) = x^2$

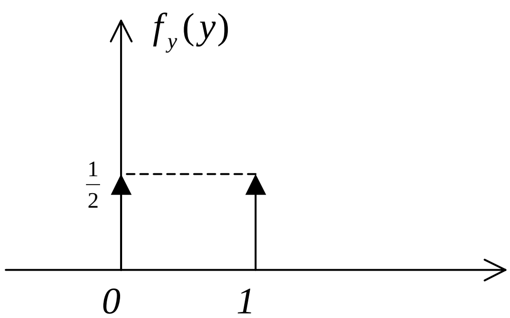
What is R_{y_0} ?

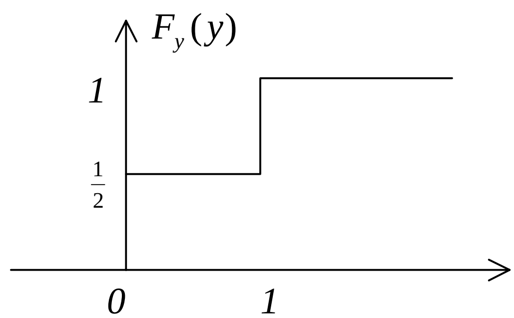


$$F_y(y_0) = \Pr(y \le y_0)$$
$$= \Pr(-\sqrt{y_0} \le x \le \sqrt{y_0})$$
$$= F_x(\sqrt{y_0}) - F_x(-\sqrt{y_0})$$

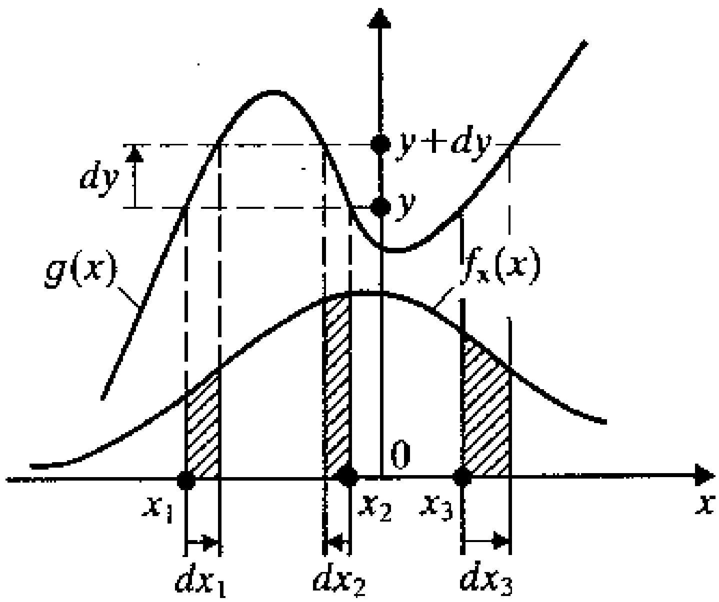
Example: Let $x \sim N(\mu, \sigma)$

And $$y = U(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \le 0 \end{cases}$$





To determine the density of $y=g(x)$ in terms of $f_x(x_0)$, look at $g(x)$



$$f_y(y_0)dy_0 = \Pr(y_0 \le y \le y_0 + dy_0)$$
$$= \Pr(x_1 \le x \le x_1 + dx_1)$$
$$+ \Pr(x_2 + dx_2 \le x \le x_2)$$
$$+ \Pr(x_3 \le x \le x_3 + dx_3)$$
$$= f_x(x_1)dx_1 + f_x(x_2)\left|dx_2\right| + f_x(x_3)dx_3$$

Note that

$$dx_1 = \frac{dx_1}{dy_0} dy_0 = \frac{dy_0}{dy_0 / dx_1} = \frac{dy_0}{g'(x_1)}$$

Similarly

$$dx_2 = \frac{dy_0}{g'(x_2)} \quad \text{and} \quad dx_3 = \frac{dy_0}{g'(x_3)}$$

Thus

$$f_y(y_0) dy_0$$

$$= \frac{f_x(x_1)}{g'(x_1)} dy_0 + \frac{f_x(x_2)}{|g'(x_2)|} dy_0 + \frac{f_x(x_3)}{g'(x_3)} dy_0$$

or

$$f_y(y_0) = \frac{f_x(x_1)}{g'(x_1)} + \frac{f_x(x_2)}{|g'(x_2)|} + \frac{f_x(x_3)}{g'(x_3)}$$

In general, for $y=g(x)$, let $x_1, x_2, \ldots$ be the roots

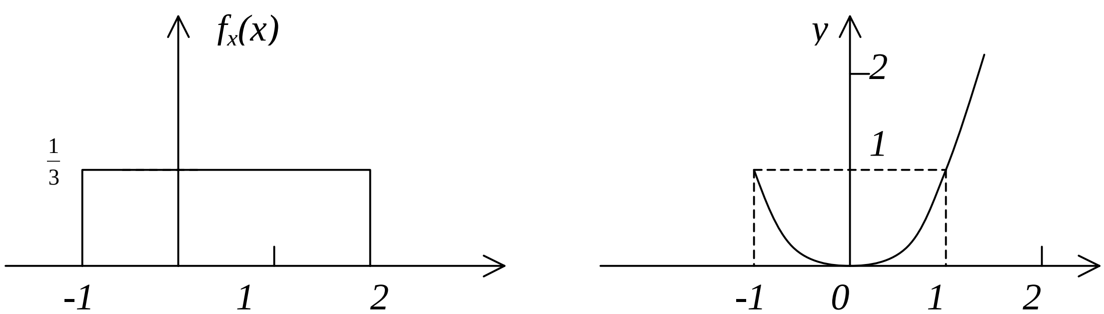
$$y=g(x_1)=g(x_2)=\ldots$$

then

$$f_y(y) = \frac{f_x(x_1)}{|g'(x_1)|} + \frac{f_x(x_2)}{|g'(x_2)|} + \ldots\ldots$$

Example: suppose $x \sim U(-1,2)$ and

$$y = x^2$$



$$g'(x) = 2x$$

case 1: $0 \le y \le 1$

$$y = x^2 \Rightarrow x = \pm\sqrt{y}$$

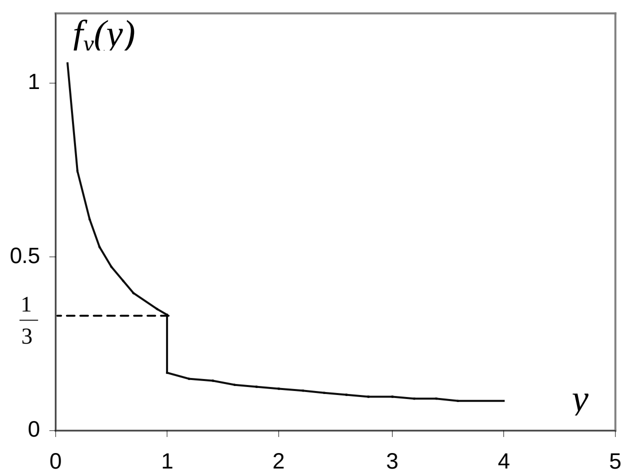
$$f_y(y) = \frac{f_x(x_1)}{\left|g'(x_1)\right|} + \frac{f_x(x_2)}{\left|g'(x_2)\right|}$$

$$= \frac{1/3}{\left|2\sqrt{y}\right|} + \frac{1/3}{\left|-2\sqrt{y}\right|} = \frac{1/3}{\sqrt{y}}$$

case 2: $1 \le y \le 4$

$$y = x^2 \Rightarrow x = \sqrt{y}$$

$$f_y(y) = \frac{f_x(x_1)}{\left|g'(x_1)\right|} = \frac{1/3}{2\sqrt{y}} = \frac{1/6}{\sqrt{y}}$$



Example: let $x \sim N(\mu, \sigma)$ and $y = e^x$

Then $g(x) \geq 0$ and $g'(x) = e^x$

Also, there is only one solution

$x = \ln(y)$

Therefore

$$f_y(y) = \frac{f_x(x)}{|g'(x)|} = \frac{f_x(x)}{e^x}$$

or

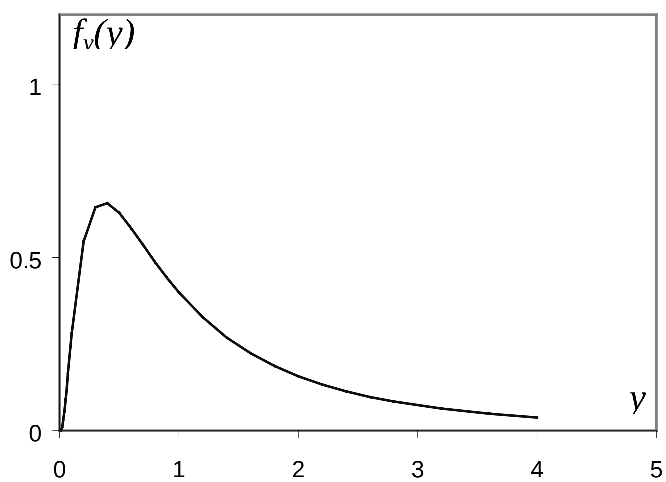
$$f_y(y) = \frac{f_x(\ln(y))}{e^{\ln(y)}} = \frac{f_x(\ln(y))}{y}$$

for

$$f_x(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

$$\Rightarrow f_y(y) = \frac{1}{\sqrt{2\pi}y\sigma} e^{\frac{-(\ln(y)-\mu)^2}{2\sigma^2}}$$

$$for \quad y > 0$$



Log normal density

Distribution of $F_x(x)$

For any RV with continuous distribution $F_x(x)$, the RV $y = F_x(x)$ is uniform on [0,1].

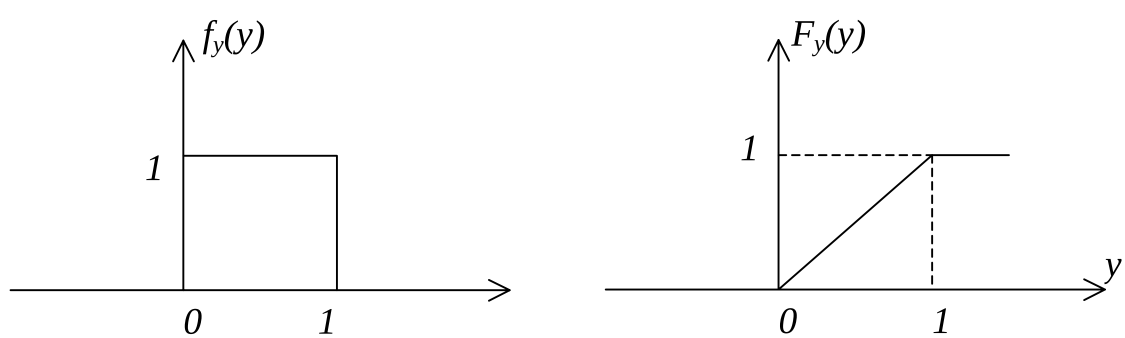
Proof: Note $0<y<1$

Since $g(x) = F_x(x)$

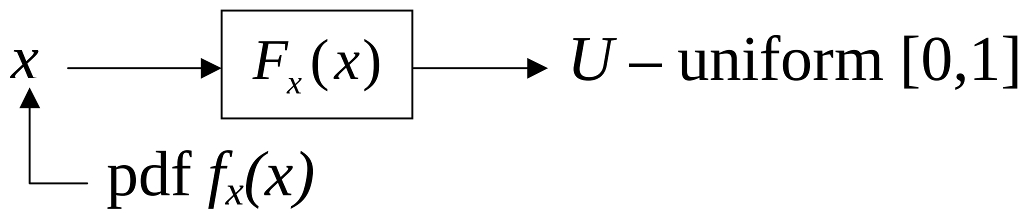
$$g'(x) = f_x(x)$$

and

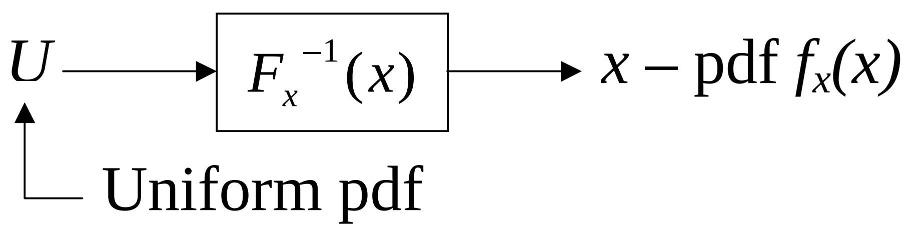
$$f_y(y) = \frac{f_x(x)}{g'(x)} = \frac{f_x(x)}{f_x(x)} = 1$$



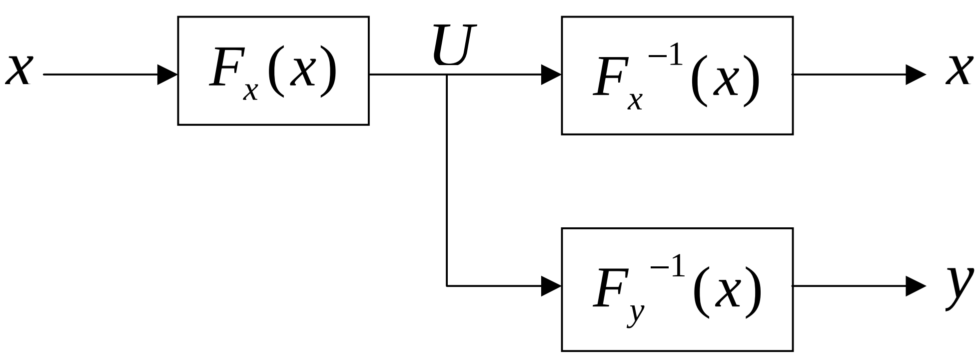
Thus the function $g(x) = F_x(x)$ performs the mapping



Similarly



Synthesis:



Mean and variance

$$E\{x\} = \int_{-\infty}^{\infty} xf(x)\,dx$$

Conditional mean

$$E\{x|M\} = \int_{-\infty}^{\infty} xf(x|M)\,dx$$

Example: If $M = \{x \geq a\}$

Then

$$E\{x|M\} = \int_{-\infty}^{\infty} xf(x|M)\,dx$$

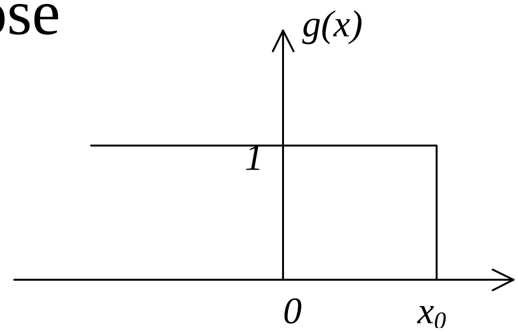
$$= \frac{\int_{a}^{\infty} xf(x)\,dx}{\int_{a}^{\infty} f(x)\,dx}$$

For a function of a RV, *y=g(x)*,

$$E\{y\} = \int_{-\infty}^{\infty} y f_y(y)\,dy$$

$$= \int_{-\infty}^{\infty} g(x)\, f_x(x)\,dx$$

Example: Suppose



Then

$$E\{g(x)\} = \int_{-\infty}^{\infty} g(x)\, f_x(x)\,dx$$

$$= \int_{-\infty}^{x_0} f_x(x)\,dx$$

$$= F_x(x_0)$$

The variance is defined as

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\eta)^2 f(x)\,dx$$

Where $\eta = E\{x\}$

Thus,

$$\sigma^2 = E\{(x-\eta)^2\} = E\{x^2\} - E^2\{x\}$$

Example: $x \sim N(\eta, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\eta)^2}{2\sigma^2}}$$

f(x) is symmetric about

$$x = \eta \Rightarrow E\{x\} = \eta$$

also

$$\int_{-\infty}^{\infty} f(x)dx = 1 \Rightarrow \int_{-\infty}^{\infty} e^{\frac{-(x-\eta)^2}{2\sigma^2}} dx = \sqrt{2\pi}\sigma$$

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

$\Downarrow$ by differentiating

$$\int_{-\infty}^{\infty} \frac{(x-\eta)^2}{\sigma^3} e^{\frac{-(x-\eta)^2}{2\sigma^2}} dx = \sqrt{2\pi}$$

$\Downarrow$ rearranging

$$\int_{-\infty}^{\infty} (x-\eta)^2 \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\eta)^2}{2\sigma^2}} dx = \sigma^2$$

or

$$E\{(x-\eta)^2\} = \sigma^2$$

Moments

(Moments)

$$m_n = E\{x^n\} = \int_{-\infty}^{\infty} x^n f(x)dx$$

(central moments)

$$\mu_n = E\{(x-\eta)^n\}$$
$$= \int_{-\infty}^{\infty} (x-\eta)^n f(x)dx$$

From the binomial theorem

$$\mu_n = E\{(x-\eta)^n\}$$
$$= E\{\sum_{k=0}^{n} \binom{n}{k} x^k (-\eta)^{n-k}\}$$
$$= \sum_{k=0}^{n} \binom{n}{k} m_k (-\eta)^{n-k}$$

$\mu_0 = 1 \qquad \mu_1 = 0 \qquad \mu_2 = \sigma^2$

$\mu_3 = m_3 = 3\eta m_2 + 2\eta^3$

Example: $x \sim N(0, \sigma^2)$

Then

$$E\{x^n\} = \begin{cases} 0 & n = 2k+1 \\ 1 \cdot 3 \cdot \cdots \cdot (n-1)\sigma^n & n = 2k \end{cases}$$

Proof: For *n* odd

$$E\{x^n\} = \int_{-\infty}^{\infty} \underbrace{x^n}_{\text{Odd function}} \underbrace{f(x)}_{\text{Even function}} dx = 0$$

To prove the second part, use the fact that

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}}$$

Differentiate both sides with respect to α,

k times

$$\int_{-\infty}^{\infty} x^{2k} e^{-\alpha x^2} dx = \frac{1\cdot 3\cdot\cdots\cdot(2k-1)}{2^k}\sqrt{\frac{\pi}{\alpha^{2k+1}}}$$

let $\alpha = \dfrac{1}{2\sigma^2}$, then

$$\int_{-\infty}^{\infty} x^{2k} e^{-\frac{x^2}{2\sigma^2}} dx = 1\cdot 3\cdot\cdots\cdot(2k-1)\sigma^{2k+1}\sqrt{2\pi}$$

Letting $n=2k$ and rearranging

$$\int_{-\infty}^{\infty} x^n \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2}{2\sigma^2}} dx = 1\cdot 3\cdot\cdots\cdot(n-1)\sigma^n$$

Variance is a measure of a RV's concentration around its mean

<u>Tchebycheff Inequality</u>

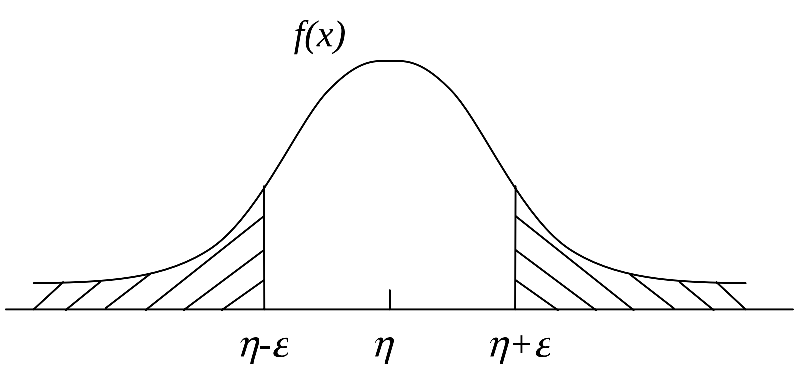
For any $\varepsilon > 0$,

$$\Pr(|x-\eta| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

Proof:

$$\Pr(|x-\eta| \geq \varepsilon) = \int_{-\infty}^{\eta-\varepsilon} f(x)dx + \int_{\eta+\varepsilon}^{\infty} f(x)dx$$

$$= \int_{|x-\eta| \geq \varepsilon} f(x)dx$$



But

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\eta)^2 f(x)\,dx$$

$$\geq \int_{|x-\eta|\geq\varepsilon} (x-\eta)^2 f(x)\,dx$$

and since $|x-\eta| \geq \varepsilon$

$$\sigma^2 \geq \int_{|x-\eta|\geq\varepsilon} \varepsilon^2 f(x)\,dx$$

$$= \varepsilon^2 \int_{|x-\eta|\geq\varepsilon} f(x)\,dx$$

$$= \varepsilon^2 \Pr\{|x-\eta| \geq \varepsilon\}$$

$$\Rightarrow \Pr\{|x-\eta| \geq \varepsilon\} \leq \left(\frac{\sigma}{\varepsilon}\right)^2$$