Ch3 Operations on one random variable-Expectation

Previously we define a random variable as a mapping from the sample space to the real line

We will now introduce some operations on the random variable. Most of these operations are based on the concept "expectation"

Expectation

Expectation is the name given to the process of averaging of a random variable X.

The followings are equivalents:

- Expectation or expected value of random variable X , which we use the notation E[X]
- The “mean value ” of random variable X
- The “statistical average” of random variable X

The following notation are equivalent $E[X] = \bar{X}$

Example:3.1-1

Expected value of a random variable

The everyday averaging procedure used in the above example carries over directly to RV

Example:
Let X be a random variable the has the following sample space values

$$S_X = \{1, 2, 3, 4\}$$

Now if the numbers are equally likely to occur or selected

$$\bar{X} = \frac{1+2+3+4}{4} = \underbrace{(1)}_{X=1} \underbrace{(\frac{1}{4})}_{P(X=1)} + \underbrace{(2)}_{X=2} \underbrace{(\frac{1}{4})}_{P(X=2)} + \underbrace{(3)}_{X=3} \underbrace{(\frac{1}{4})}_{P(X=3)} + \underbrace{(4)}_{X=4} \underbrace{(\frac{1}{4})}_{P(X=4)}$$

In general then

$$\bar{X} = \sum_{x_i \in S_X} x_i \; P(X = x_i)$$

where each value of the random variable X (x_i) is weighted by the Probability $P(X= x_i)$

This motivate the concept of expected value or mean value of RV X,

$$E[X] = \bar{X} = \sum_{x_i \in S_X} x_i P(x_i)$$ if X is discrete values

$$E[X] = \bar{X} = \int_{-\infty}^{\infty} x f_X(x) dx$$ if X is continuous value with Probability density

In the discrete random variable we use the probability mass $P(X = x_i)$ to weight the random variable.

In the continuous random variable we use the density $f_X(x)$ to weight the random variable.

Observe that $f_X(x)dx$ represent the probability of the random variable X at the interval dx

If the random variable X is symmetrical about a line $x = a$

$$\Rightarrow f_X(x + a) = f_X(-x + a) \qquad \Rightarrow E[X] = a$$

Example 3.1-2

Expected value of a Function of a random variable

Assume a random variable X which has the following values and probabilities

$$X = \{1, 2, 3\} \qquad P(X=1) = \frac{1}{3} \qquad P(X=2) = \frac{2}{5} \qquad P(X=3) = \frac{4}{15}$$

Now define the random variable $Y = X^2$

$$\Rightarrow Y = \{1, 4, 9\} \quad P(Y=1) = \frac{1}{3} \quad P(Y=4) = \frac{2}{5} \quad P(Y=9) = \frac{4}{15}$$

$$\Rightarrow E[Y] = \underbrace{(1)}_{Y=1}\underbrace{\underbrace{\left(\frac{1}{3}\right)}_{P(Y=1)}}_{P(X=1)} + \underbrace{(4)}_{Y=4}\underbrace{\underbrace{\left(\frac{2}{5}\right)}_{P(Y=4)}}_{P(X=2)} + \underbrace{(9)}_{Y=9}\underbrace{\underbrace{\left(\frac{4}{15}\right)}_{P(Y=9)}}_{P(X=3)} = \sum_{y_i \in \{1,4,9\}} y_i P(Y=y_i)$$

$$= \sum_{y_i \in \{1,4,9\}} y_i P(X=x_i)$$

In general then

$$E\left[g(X)\right] = \sum_{i=1}^{N} g(x_i)P(x_i)$$

for discrete random variable

$$E\left[g(x)\right] = \int_{-\infty}^{\infty} g(x)f_X(x)dx$$

for continuous random variable

Conditional Expectation

We define the conditional density function for a given event

$B = \{ X \le b\}$

$$f_X(x|X \le b) = \begin{cases} \dfrac{f_X(x)}{\int_{-\infty}^{b} f_X(x)dx} & x < b \\ 0 & x \ge b \end{cases}$$

we now define the conditional expectation in similar manner

$$E[X|B] = \int_{-\infty}^{\infty} xf_X(x|B)dx = \underbrace{\int_{-\infty}^{b} xf_X(x|B)dx}_{x < b} + \underbrace{\int_{b}^{\infty} xf_X(x|B)dx}_{x \ge b}$$

$$= \int_{-\infty}^{b} x \frac{f_X(x)}{\underbrace{\int_{-\infty}^{b} f_X(x)dx}_{\text{constant} = F_X(b)}} dx + \underbrace{\int_{b}^{\infty} x\, 0\, dx}_{0} = \frac{\int_{-\infty}^{b} xf_X(x)dx}{\underbrace{\int_{-\infty}^{b} f_X(x)dx}_{\text{constant} = F_X(b)}}$$

Moments

The expected value defined previously as

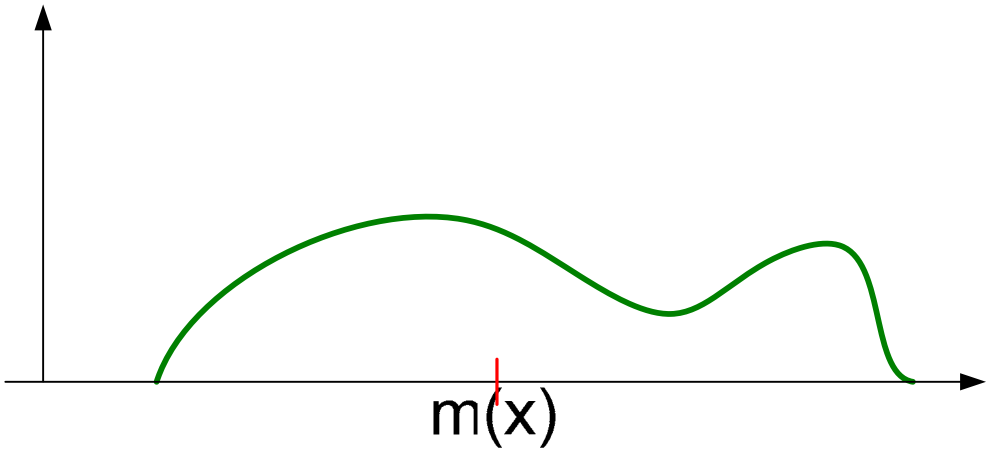
$$E[X] = \bar{X} = \int_{-\infty}^{\infty} x f_X(x) dx$$

if X is continuous value with probability density $f_X(x)$

Is called the 1st ***moment*** of the random variable X with probability density $f_X(x)$

The word ***moment*** is used because a similar form exist in static were the 1st ***moment*** represent the ***center of gravity***

Example Assume a mass is distributed on one dimension x as shown below were m(x) is the mass density function is



Then we can calculate the 1st ***moment*** or ***center of gravity*** M as

$$M = \frac{\int_a^b x\, m(x)\, dx}{\underbrace{\int_a^b m(x)\, dx}_{\text{total mass}}}$$

In the probability case the total area under the density function is unity or 1

The expected value of the function of random variable X was defined as

$$E[g(x)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

Let the function $g(X)$ defined as

$$g(X) = X^n \qquad n = 0, 1, 2, \ldots$$

Then we can define the n^{th} moment m_n as

$$m_n = E\left[X^n\right] = \int_{-\infty}^{\infty} x^n f_X(x) dx$$

Clearly

$$m_0 = \underbrace{E\left[X^0\right]}_{E[1]=1} = \underbrace{\int_{-\infty}^{\infty} x^0 f_X(x)dx}_{\int_{-\infty}^{\infty} f_X(x)dx = 1} = 1$$ The area of the function $f_X(x)$

$$m_1 = E\left[X^1\right] = \int_{-\infty}^{\infty} x f_X(x)dx = \bar{X}$$ The expected value of X

Central Moments

Another type of moments of interest is the central moment defined as

$$\mu_n = E\left[(X - \bar{X})^n\right] = \int_{-\infty}^{\infty}(x - \bar{X})^n f_X(x)dx$$

the m_n moments are expected values about the origin however the central moment m_n is moment or expected value about the mean or average $\bar{X}$

$$\mu_0 = \underbrace{E\left[(X-\bar{X})^0\right]}_{E[1]=1} = \underbrace{\int_{-\infty}^{\infty}(X-\bar{X})^0 f_X(x)dx}_{\int_{-\infty}^{\infty} f_X(x)dx = 1} = 1$$

The area of the function $f_X(x)$

$$\mu_1 = E\left[(X-\bar{X})^1\right] = E[X] - E[\bar{X}] = \bar{X} - \bar{X} = 0$$

were we have used the fact that $E[\underbrace{a}_{\text{constant}}] = a$

Variance and skew

The second central moment μ_2 is so important , that it is given the name ***variance*** and have the special notation σ_x^2

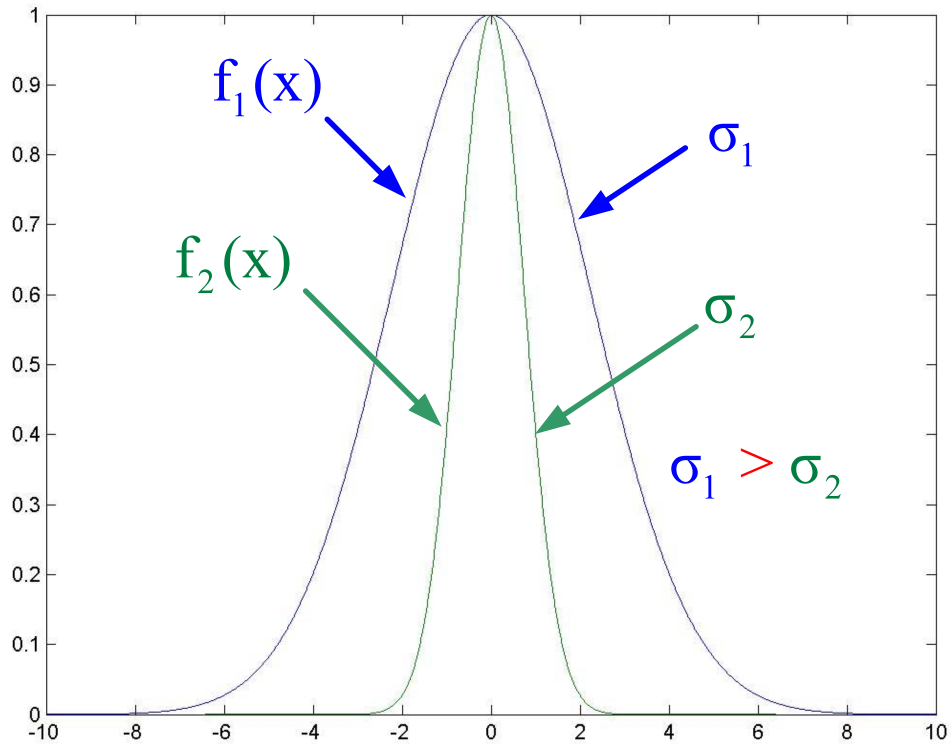
Thus the variance is given by

$$\sigma_x^2 = \mu_2 = E\left[(X - \bar{X})^2\right] = \int_{-\infty}^{\infty} (x - \bar{X})^2 f_X(x)dx$$

σ_X (the positive square root of the variance) is called the standered deviation of RV X

σ_X is a measure of the spread of the random variable X about its mean or average

σ_X is a measure of the spread of the random variable X about its mean or average



The spread of $f_1(x)$ is more than the spread of $f_2(x)$ $(\sigma_1 > \sigma_2)$

variance can be found from a knowledge of first and second moments as follows

$$\sigma_x^2 = \mu_2 = E\left[X^2 - 2\overline{X}X + \overline{X}^2\right] = E\left[X^2\right] - 2\overline{X}E[X] + \overline{X}^2$$

$$= E\left[X^2\right] - \overline{X}^2 = m_2 - m_1^2$$

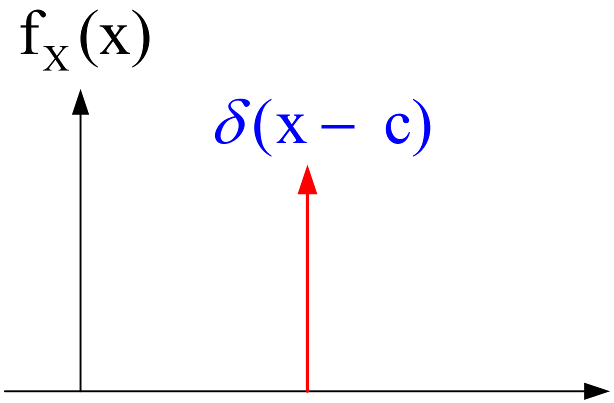
Example 3.2-1

Properties of the variance σ_X

Let c be a constant and X be a RV

(1) σ_c^2

The probability density function of a constant "deterministic number" is a delta function with spread zero



(2) $\sigma^2_{x+c} = \sigma^2_x$

The variance does not change by shifting the random variable, the spread will remain the same , on the other hand shifting effect the mean only

(3) $\sigma^2_{cx} = c^2\sigma^2_x$

The third central moment $\mu_3 = E\left[(X - \bar{X})^3\right]$ is a measure of the a symmetry of $f_X(x)$ about $x = \bar{X} = m_1$

It will be called the "skew" of the density function

If $f_X(x)$ symmetric about $x = \bar{X}$, it has zero skew (i.e $\mu_3 = 0$)

$\mu_n = 0$ for all odd n .

The normalized third central moment μ_3/σ^3_x is known as the skewness or coefficient of skewness of the probability density function

Useful Inequalities

A useful tool in some probability problems are some inequalities such as Chebychev's inequality and Markov's Inequality.

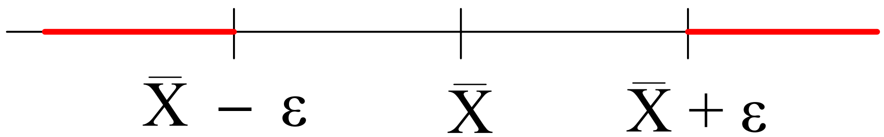
Chebychev's Inequality

It state that for a RV X ,

$$P\left\{\left|X - \overline{X}\right| \geq \epsilon\right\} \leq \sigma_X^2/\epsilon^2 \qquad \text{for any } \varepsilon > 0$$

$$P\left\{\left|X - \bar{X}\right| \geq \epsilon\right\} \leq \sigma_X^2 / \epsilon^2$$

Proof



$$P\left\{\left|X - \bar{X}\right| \geq \epsilon\right\} = P\left\{X - \bar{X} \geq \epsilon \ \text{ and } \ X - \bar{X} \leq -\epsilon\right\}$$

$$= P\left\{X \geq \bar{X} + \epsilon \quad \text{and} \quad X \leq \bar{X} - \epsilon\right\}$$

$$= \int_{-\infty}^{\bar{X}-\epsilon} f_X(x)dx + \int_{\bar{X}+\epsilon}^{\infty} f_X(x)dx = \int_{\left|x - \bar{X}\right| \geq \epsilon}^{\infty} f_X(x)dx$$

but since $\sigma_X^2 = \int_{-\infty}^{\infty} (x - \bar{X})^2 f_X(x)dx \geq \int_{\left|x - \bar{X}\right| \geq \in}^{\infty} \underbrace{(x - \bar{X})^2}_{(x - \bar{X})^2 \geq \varepsilon^2} f_X(x)dx$

$$\geq \in^2 \int_{\left|x - \bar{X}\right| \geq \in}^{\infty} f_X(x)dx = \in^2 P\left\{\left|X - \bar{X}\right| \geq \in\right\}$$

$$\Rightarrow P\left\{\left|X - \bar{X}\right| \geq \epsilon\right\} \leq \sigma_X^2 / \epsilon^2$$

$$P\{|X - \bar{X}| \geq \epsilon\} \leq \sigma_X^2/\epsilon^2$$

Another form of Chebychev's Inequality is

$$P\{|X - \bar{X}| < \in\} \geq 1 - (\sigma_X^2/\in^2)$$

A consequence of that if $\sigma_X^2 \to 0$ for a random variable , then

$$P\{|X - \bar{X}| < \in\} \to 1 \quad \text{or} \quad P\{X = \bar{X}\} \to 1$$

In other words, if the variance of the RV X approach zero, the probability approaches 1 , that X will equal its mean.

Markov's Inequality

Let X be a non negative random number , then

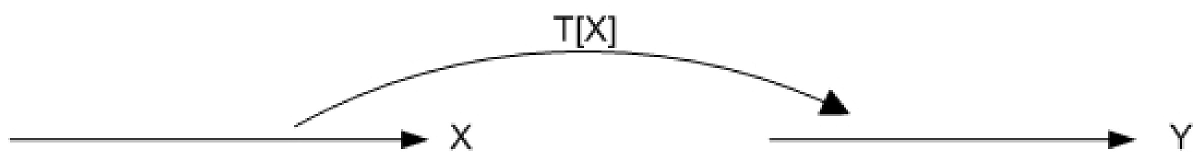
$$P\{X \geq a\} \leq E[X]/a \quad - a > 0$$

3.4 Transformations of A Random Variable

Let X be a random variable with a known distribution $F_X(x)$ and a known density $f_X(x)$.

Let T(.) be a transformation or a mapping or a function that maps the R.V X into Y as

$$Y = T(X)$$

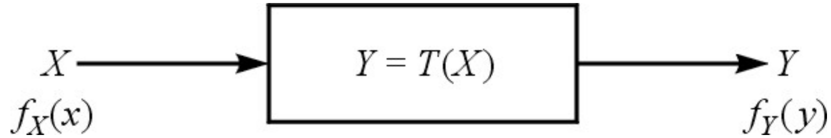


$$F_X(x) = P[X \le x] \qquad F_Y(y) = P[Y \le y] \quad ?$$

$$f_X(x) = \frac{dF_X(x)}{dx} \qquad f_Y(y) = \frac{dF_Y(y)}{dy} \quad ?$$

The problem is to find $F_Y(y)$ and $f_Y(y)$.

You can view the problem as a block box problem



In general X can be a discrete, continuous or mixed random variable and the Transformation T can be

Linear

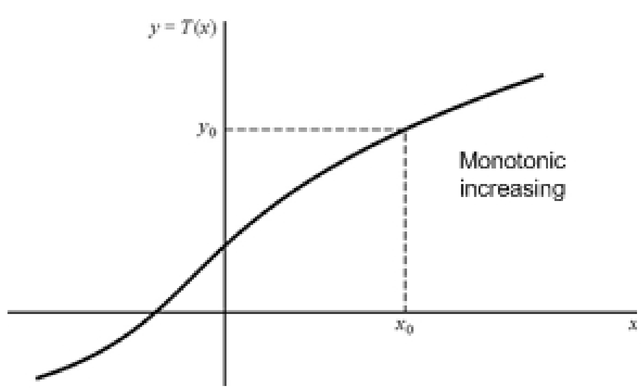
Non-linear

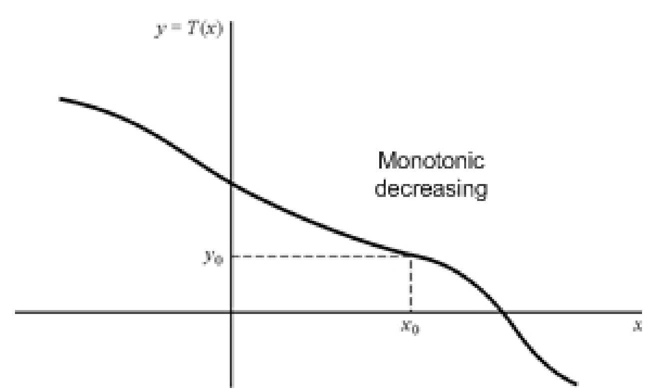
Segmented

Staircase

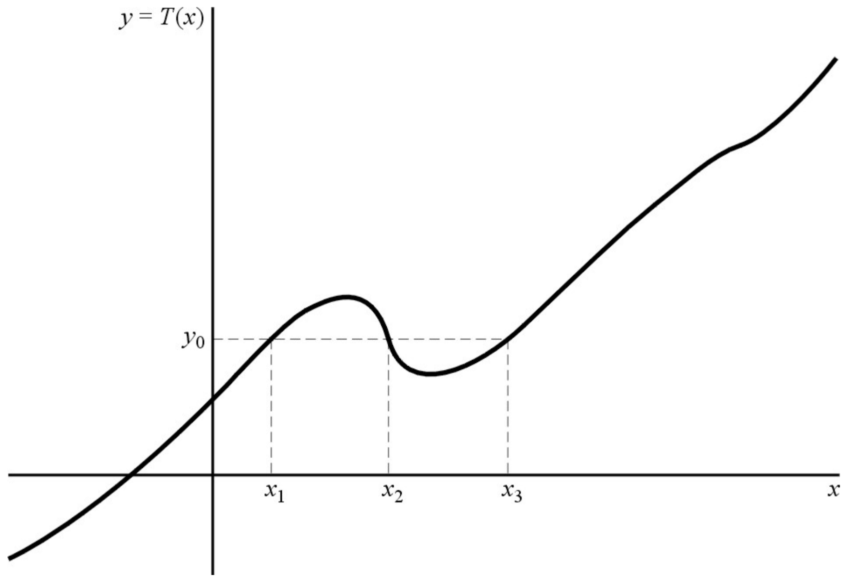
We will consider three cases:

(1) X is continuous and T is continuous and monotonically increasing or decreasing





(2) X is continuous and T is continuous nonmonotonic

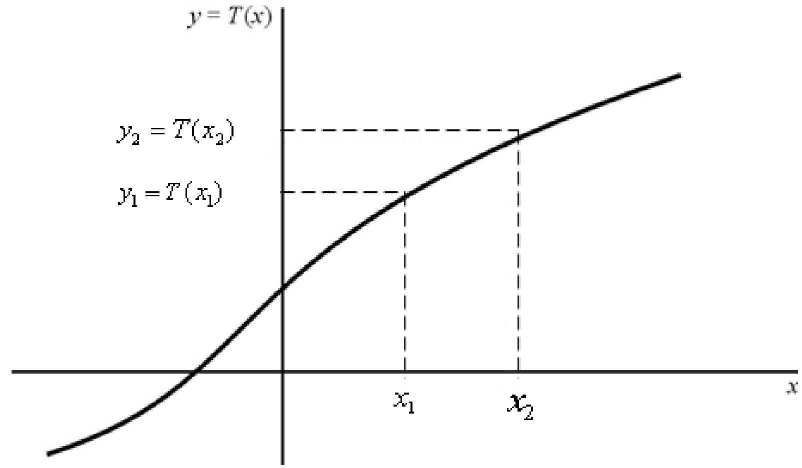


(3) X is discrete and T is continuous

Monotonic Increasing Transformations of a Continuous Random Variable

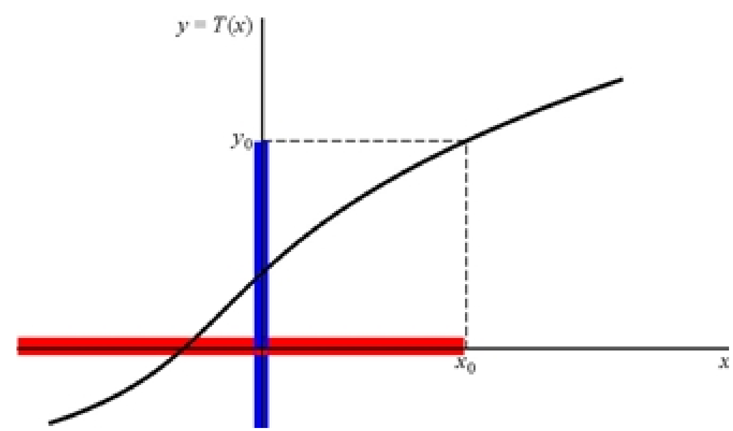
A transformation T is called monotonically increasing if

$T(x_1) < T(x_2)$ for any $x_1 < x_2$.

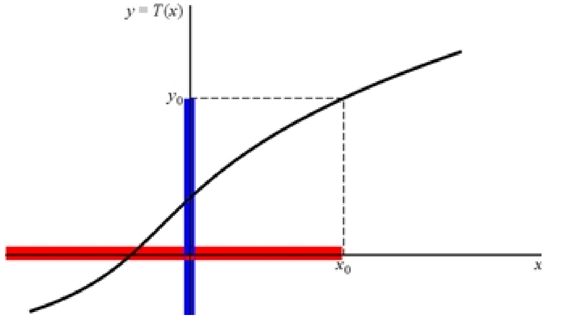


Assume that T is continuous and differentiable at all values of x for which $f_X(x) \neq 0$.

$$y_0 = T(x_0) \; or \; x_0 = T^{-1}(y_0)$$



The events $\{Y \leq y_0\}$ and $\{X \leq x_0\}$ are equivalent.



$\Rightarrow$ P $\{Y \leq y_0\}$ = P $\{X \leq x_0\}$

$$\Rightarrow F_Y(y_0) = P\{Y \leq y_0\} = P\{X \leq x_0\} = F_X(x_0)$$

$$\Rightarrow \int_{-\infty}^{y_0} f_Y(y)dy = \int_{-\infty}^{x_0 = T^{-1}(y_0)} f_X(x)dx$$

Differentiating both sides with respect to y_0 and using Leibniz rule

Let

$$G(u) = \int_{\alpha(u)}^{\beta(u)} H(x,u)dx$$

$$\frac{dG(u)}{du} = H[\beta(u),u]\frac{d\beta(u)}{du} - H[\alpha(u),u]\frac{d\alpha(u)}{du} + \int_{\alpha(u)}^{\beta(u)} \frac{\partial H(x,u)}{\partial u}dx$$

Now the LHS,

$$\frac{d}{dy_0}\int_{-\infty}^{y_0} f_Y(y)dy = f_Y(y_0)\frac{dy_0}{dy_0} - f_Y(y_0)(0) + \int_{-\infty}^{y_0}\frac{df_Y(y)}{dy_0}dy$$

$$= f_Y(y_0)$$

The RHS

$$\frac{d}{dy_o}\int_{-\infty}^{x_0=T^{-1}[y_0]} f_X(x)dx = f_X[x_0]\frac{dx_0}{dy_0} - f_X(x_0)(0) + \int_{-\infty}^{x_0=T^{-1}[y_0]}\frac{df_X(x)}{dy_0}dx$$

$$= f_X(T^{-1}(y_0))\frac{dT^{-1}(y_0)}{dy_0}$$

$$\Rightarrow f_Y(y_0) = f_X(T^{-1}(y_0))\frac{dT^{-1}(y_0)}{dy_0}$$

Since the results apply for any y_0,

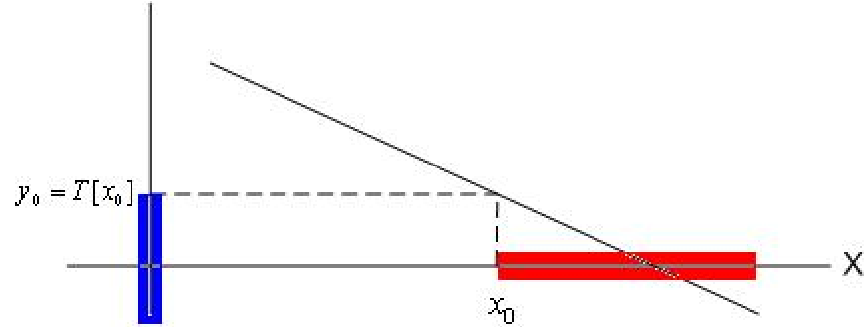
$$\Rightarrow f_Y(y) = f_X\left[T^{-1}(y)\right] \frac{dT^{-1}(y)}{dy}$$

$$or \quad f_Y(y) = f_X(x) \frac{dx}{dy}$$

Monotonic Decreasing Transformations of a Continuous Random Variable

A transformation T is called monotonically decreasing if

$T(x_1) > T(x_2)$ for any $x_1 < x_2$.



$$F_Y(y_0) = P\{Y \le y_0\} = P\{X \ge x_0\} = 1 - F_X(x_0)$$

$$\int_{-\infty}^{y_0} f_Y(y)dy = 1 - \int_{-\infty}^{x_0 = T^{-1}(y_0)} f_X(x)dx$$

$$\frac{d}{dy_0}\int_{-\infty}^{y_0} f_Y(y)dy = f_Y(y) = 0 - f_X[T^{-1}(y_0)]\frac{dT^{-1}(y_0)}{dy_0}$$

since $\dfrac{dT^{-1}(y_0)}{dy_0}$ is negative (monotone decreasing)

$$\Rightarrow f_Y(y) = f_X\left[T^{-1}(y)\right] \left|\frac{dT^{-1}(y)}{dy}\right| = f_X(x)\left|\frac{dx}{dy}\right|$$

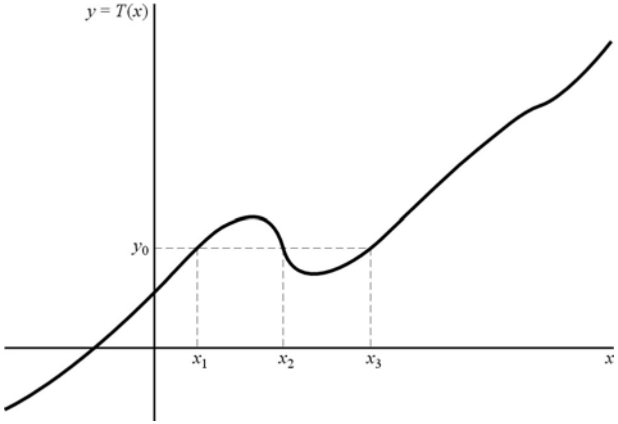
Then we conclude that

$$f_Y(y) = f_X(x) \left|\frac{dx}{dy}\right|$$

for both increasing and decreasing monotonic transformation.

Nonmonotonic Transformations of a Continuous Random Variable

A transformation may not be monotonic(increasing and decreasing) in the more general case it can be both increasing and decreasing as shown below



The event $\{ \ Y \leq y_0 \ \}$ may correspond to more than one event of the random variable X.