Math 3501 - Probability and Statistics I

5.1 - Functions of one random variable

Discrete random variables

Let X be a discrete random variable with pmf

$$f(x) = P(X = x) \ , \quad x \in S_X = \{c_1, c_2, c_3, \ldots\}$$

and let $u(x)$ be a real-valued function of a single real variable.

Define a new random variable

new old r.v

$$\boxed{Y = u(X) \ .}$$

and note that Y is a discrete random variable with support

$$S_Y = \{d_1, d_2, d_3, \ldots\} \ , \quad \text{where } d_i = u(c_i) \text{ for each } i = 1, 2, \ldots \ .$$

The values d_i are not necessarily distinct if u is not one-to-one.

S_x : c_1 c_2 c_3 c_4 ...

S_y d_1 d_2 d_3

$d_1 = u(c_1)$, $d_2 = u(c_2)$, $d_3 = u(c_3)$

The pmf of $Y = u(X)$ is given by

$$g(y) = P(Y = y) = P[u(X) = y] = \sum_{\{x: u(x) = y\}} f(x) \qquad y \in S_Y .$$

Note: for each $y \in S_Y$, the value of $g(y)$ is found by summing probabilities over all values of x for which $u(x)$ equals y.

v is u^{-1}

Special case: If $y = u(x)$ is a one-to-one transformation with inverse $x = v(y)$, the pmf of $Y = u(X)$ is given by

$$g(y) = f(v(y)) , \quad y \in S_Y .$$

u one-to-one

$$g(y) = P(Y = y) = P(u(X) = y) = P(X = u^{-1}(y)) = P(X = v(y)) = f(v(y))$$

Example

Suppose X has a Poisson distribution with mean $\lambda = 4$.

Find the pmf of $Y = \sqrt{X}$.

space X = support of X
= support of pmf $f(x)$

$X \sim$ Poisson (4) means that the pmf of X is

$$\longrightarrow \quad f(x) = \frac{4^x e^{-4}}{x!}, \quad x = \underbrace{\overbrace{0, 1, 2, 3, \ldots}}$$

$S_x = \{0, 1, 2, 3, \ldots\}$

Since $Y = \sqrt{X}$ then $S_y = \{0, 1, \sqrt{2}, \sqrt{3}, \ldots\}$

If we let $Y = u(X)$ with $u(x) = \sqrt{x}$,

we observe that $u(x) = \sqrt{x}$ is 1-to-1 on the set $S_x = \{0, 1, 2, 3, \ldots\}$

Thus, the pmf of Y is $g(y) = P\{Y = y\} = P(\sqrt{X} = y) = P(X = y^2) = f(y^2) = \dfrac{4^{y^2} e^{-4}}{(y^2)!}$,

for $y = 0, 1, \sqrt{2}, \sqrt{3}, \ldots$

Continuous random variables: cdf technique

Let X be a random variable of the continuous type.

Define the random variable

$$\boxed{Y = u(X)} \ ,$$

where u is a given real-valued function of a single real variable.

The cdf of Y is given by

by def of cdf   by def of Y

$$G(y) = P(Y \leq y) = P[u(X) \leq y] \ . \quad = \ ???$$

If Y is also a random variable of the continuous type, then its pdf is given by

$$\boxed{g(y) = G'(y)}$$

pdf = derivative of cd

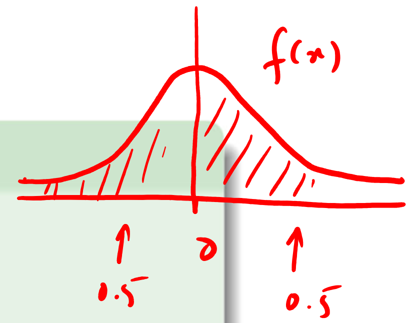
for every value $y \in \mathbb{R}$ where G is differentiable.

Note: We employed a similar strategy while deriving the cdf and pdf of the exponential distribution from the pmf of the Poisson distribution.

Remark: X being a random variable of continuous type, does NOT necessarily imply that

$$Y = u(X)$$

is also a random variable of the continuous type.



Example

X is continuous

Let $X \sim N(0,1)$ and

$$u(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}.$$

Then $Y = u(X) \sim \text{Bernoulli}(0.5)$ is a discrete random variable despite X being continuous.

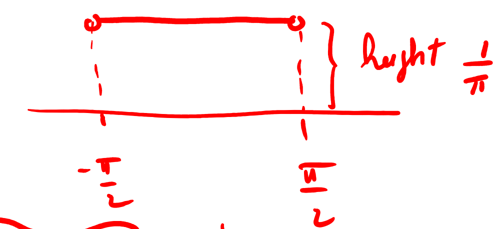
Y takes only two values $\begin{cases} 1 & \text{if } X \geq 0 \\ 0 & \text{if } X < 0 \end{cases} \Rightarrow$ Y is discrete

Example

If $X \sim U\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, find the distribution of the random variable $Y = \tan(X)$.

If $X \sim \text{Uniform}\underbrace{\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)}$, then X has pdf $f(x) = \begin{cases} \frac{1}{\pi}, & x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \\ 0, & \text{otherwise} \end{cases}$

length of interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ is $\frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi$

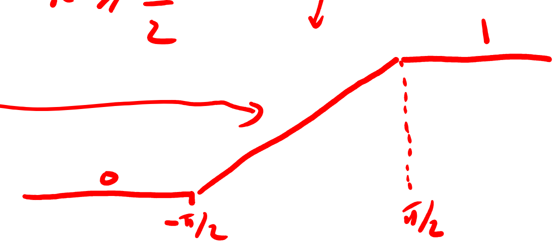


The cdf of X is then:

From Sec 3.1

$$F(x) = \begin{cases} 0, & \text{if } x \leq -\frac{\pi}{2} \\ \int_{-\pi/2}^{x} \frac{1}{\pi}\, du & \text{if } -\frac{\pi}{2} < x < \frac{\pi}{2} \\ 1, & \text{if } x \geq \frac{\pi}{2} \end{cases} = \begin{cases} 0 & \text{if } x \leq -\frac{\pi}{2} \\ \frac{x + \pi/2}{\pi} & \text{if } \boxed{-\frac{\pi}{2} < x < \frac{\pi}{2}} \\ 1 & \text{if } x \geq \frac{\pi}{2} \end{cases}$$

$\frac{x + \pi/2}{\pi}$



To find the distribution of Y:

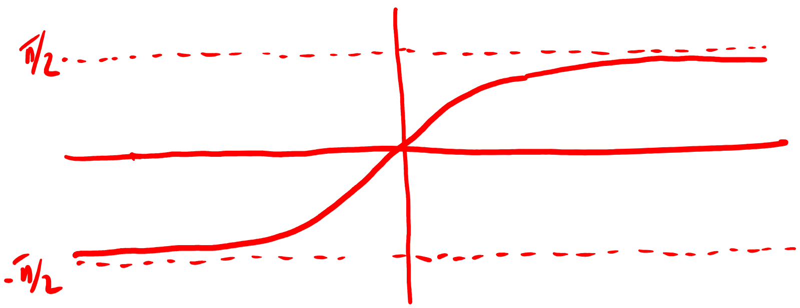
$$G(y) \underset{\text{def of cdf}}{=} P(Y \le y) \underset{\text{def of } Y}{=} P(\tan(X) \le y) \le P(X \le \arctan(y))$$

def of cdf of X

$$= F(\arctan(y))$$

$$= \frac{\arctan(y) + \pi/2}{\pi}$$

substitute $x = \arctan(y)$ in $\frac{x + \pi/2}{\pi}$ because $\arctan(y) \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$



Range of arctan is $(-\pi/2, \pi/2)$

CONCLUSION: The cdf of Y is

$$\boxed{G(y) = \frac{\arctan(y) + \pi/2}{\pi}, \quad y \in \mathbb{R}}$$

The pdf is

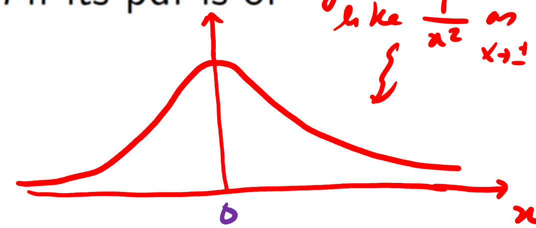
$$\boxed{g(y) = G'(y) = \frac{1}{\pi} \frac{1}{1+y^2}, \quad y \in \mathbb{R}}$$

Cauchy distribution

A continuous random variable is said to have a *Cauchy distribution* if its pdf is of the form

$$\boxed{g(x) = \frac{1}{\pi\,(1+x^2)}\,, \quad x \in \mathbb{R}\,.}$$

goes to zero like $\frac{1}{x^2}$ as $x \to \pm\infty$



Note: If X has a Cauchy distribution, then $E(X)$ is undefined.

shape is somewhat similar to that of the $\underbrace{\text{Gaussian or bell curve}}_{\text{pdf of normal distribution}}$

BUT

it goes to zero MUCH MUCH SLOWER when $x \to \pm\infty$

goes to zero fast!! $\rightarrow$ $e^{-x^2/2}$

KEY FEATURE:

$E[X]$ (or even $E[X^n]$ for any $n \in \mathbb{N}$) DOES NOT EXIST!!!

$g(x) = \frac{1}{\pi(1+x^2)}$, $x \in \mathbb{R}$ is a pdf because:

(1) $g(x) > 0$ for all $x \in \mathbb{R}$.

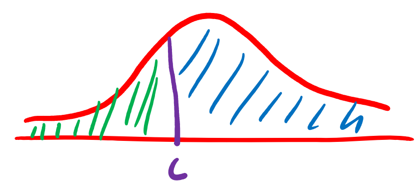
(2) $\int_{-\infty}^{\infty} g(x)\,dx = \int_{-\infty}^{\infty} \frac{1}{\pi(1+x^2)}\,dx = 1$

Reason why $E[X]$ does not exist:

By definition $E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx = \int_{-\infty}^{\infty} x \cdot \frac{1}{\pi(1+x^2)}\,dx = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x}{1+x^2}\,dx$

The improper integral is defined if and only if

$$\underbrace{\int_{-\infty}^{c} \frac{x}{1+x^2}\,dx} \quad \text{and} \quad \underbrace{\int_{c}^{\infty} \frac{x}{1+x^2}\,dx} \quad \text{both exist for any } c \in \mathbb{R}$$



and their sum does not depend on c.

For instance, if we take $c=0$, we can show that

$$\int_0^{\infty} \frac{x}{1+x^2}\,dx = \lim_{t\to\infty} \int_0^{t} \frac{x}{1+x^2}\,dx = \lim_{t\to\infty} \left[\frac{1}{2}\ln(1+x^2)\right]_{x=0}^{x=t} = \lim_{t\to\infty} \frac{1}{2}\ln(1+t^2) = \infty$$

CONCLUSION: $\int_0^{\infty} \frac{x}{1+x^2}\,dx$ is divergent $\Rightarrow \int_{-\infty}^{\infty} \frac{x}{1+x^2}\,dx$ is divergent $\Rightarrow E[X]$ does not exist

To show that $\displaystyle\int_{-\infty}^{\infty} \frac{1}{\pi(1+x^2)}\,dx = 1$:

Take $c \in \mathbb{R}$ arbitrary

$$\int_{-\infty}^{c} \frac{1}{\pi(1+x^2)}\,dx = \lim_{t\to-\infty} \int_{t}^{c} \frac{1}{\pi(1+x^2)}\,dx = \lim_{t\to-\infty} \left[\frac{1}{\pi}\arctan(x)\right]_{x=t}^{x=c}$$

$$= \lim_{t\to-\infty} \left[\frac{1}{\pi}\arctan(c) - \frac{1}{\pi}\arctan(t)\right]$$

$$= \frac{1}{\pi}\arctan c - \frac{1}{\pi}\left(-\frac{\pi}{2}\right) =$$

$$= \frac{1}{2} + \frac{1}{\pi}\arctan(c)$$

$$\int_{c}^{\infty} \frac{1}{\pi(1+x^2)}\,dx = \lim_{t\to\infty} \int_{c}^{t} \frac{1}{\pi(1+x^2)}\,dx = \lim_{t\to\infty} \left[\frac{1}{\pi}\arctan x\right]_{c}^{t} = \frac{1}{2} - \frac{1}{\pi}\arctan(c)$$

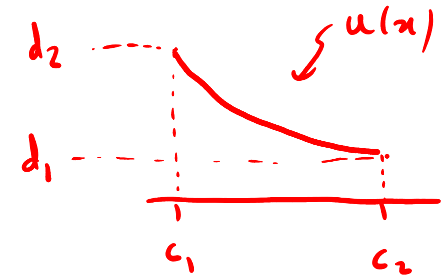
Since $\displaystyle\int_{-\infty}^{c} \frac{1}{\pi(1+x^2)}\,dx + \int_{c}^{\infty} \frac{1}{\pi(1+x^2)}\,dx = \frac{1}{2} + \frac{1}{2} = 1$ does not depend on c $\Rightarrow \displaystyle\int_{-\infty}^{\infty} \frac{1}{\pi(1+x^2)}\,dx = 1$

Change-of-variable technique

Let X be a continuous random variable with pdf $f(x)$ with support (c_1, c_2).

Let u be a real-valued strictly monotone differentiable function on (c_1, c_2) with range (d_1, d_2), and define

$$Y = u(X) \ .$$



The cdf of Y, given by

$$\boxed{G(y) = P(Y \le y) = P[u(X) \le y] \ ,}$$

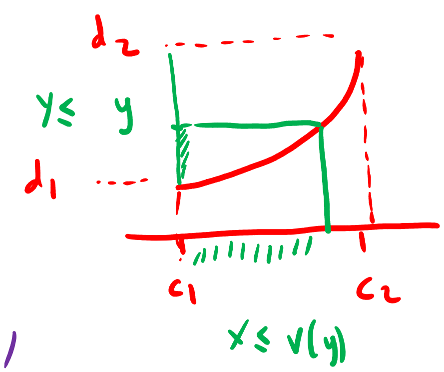
is differentiable on (c_1, c_2), and the pdf of Y is given by

$$g(y) = \begin{cases} f(v(y))\,|v'(y)| \ , & y \in (d_1, d_2) \\ 0 \ , & \text{elsewhere} \end{cases}$$

where v is the inverse of u.

DO NOT FORGET ABSOLUTE VALUE !!!

or



Proof (case where $u(x)$ is increasing) $\leftarrow$ tells us also how to proceed in the case where u is not 1-to-1

$Y = u(X)$ has cdf:

u strictly monotone $\Rightarrow$ u is 1-to-1 $\Rightarrow$ u has inverse v

$$G(y) = P(Y \leq y) = P(u(X) \leq y) = P(X \leq \underbrace{v}_{u^{-1}}(y)) = \int_{c_1}^{v(y)} \underbrace{f(x)}_{\text{pdf of } X}\, dx$$

Since $v(y)$ is differentiable, we can use the Fundamental Theorem of Calculus to compute the derivative of G:

because $v'(y) > 0$ because v is increasing

$$g(y) = G'(y) = f(v(y)) \cdot v'(y) = \underbrace{f(v(y))}_{\text{like discrete case}} \cdot |v'(y)|$$

additional term from FTC

Example

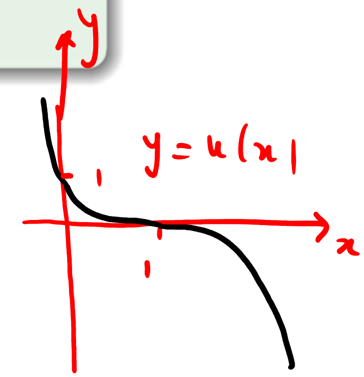
Let X have the pdf

$$f(x) = 3(1-x)^2\ , \quad \underline{x \in (0,1)}\ .$$

support of X

Find the pdf of $Y = (1-X)^3$.

Y is of the form $Y = u(X)$ with $u(x) = (1-x)^3$



u is strictly decreasing on $\underbrace{(0,1)}_{\text{support of } X}$ and so u has an inverse:

To find the inverse of u:

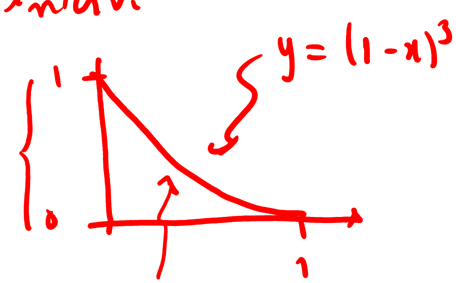
solve $y = u(x)$ for x to get $x = v(y)$.

$$y = (1-x)^3 \iff y^{1/3} = 1-x \iff x = 1 - y^{1/3}$$

The inverse is $x = v(y)$ where $v(y) = 1 - y^{1/3}$

Zoom in into the interval

range of u is (0,1)



The change of variable formula says that the pdf of $Y = u(X)$ is given by

range of u

$$g(y) = \begin{cases} f(v(y)) \cdot |v'(y)| , & y \in (0,1) \\ 0 , & \text{otherwise.} \end{cases}$$

Note that $f(v(y)) \cdot |v'(y)| = 3(1 - v(y))^2 \cdot |v'(y)| =$

$$= 3\Big(1 - \underbrace{(1 - y^{1/3})}_{v(y)}\Big)^2 \left| -\frac{1}{3} y^{-2/3} \right|$$

$$= 3\left(y^{1/3}\right)^2 \cdot \frac{1}{3} y^{-2/3} = y^{2/3} \cdot y^{-2/3} = 1$$

$$g(y) = \begin{cases} 1 & \text{if } y \in (0,1) \\ 0 & \text{otherwise} \end{cases} \Rightarrow Y \sim \text{Uniform}(0,1)$$

Remark

The change-of-variable technique extends to the case where the transformation u is not one-to-one.

Example

Find the distribution of $Y = X^2$, where X is Cauchy distributed.

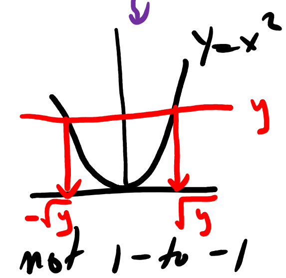
If X is Cauchy, then the pdf of X is $f(x) = \dfrac{1}{\pi(1+x^2)}$, $x \in \mathbb{R}$.

The cdf of $Y = X^2$ is

Since $Y = X^2$ and $S_X = \mathbb{R}$, then $\underbrace{S_Y = [0,\infty)}_{\text{range}}$

$$G(y) = P(Y \le y) = P(X^2 \le y) = P(-\sqrt{y} \le X \le \sqrt{y}) =$$

$u(x) = x^2$ is <u>NOT</u> 1-to-1 on $\mathbb{R}$



Then

$$G(y) = \int_{-\sqrt{y}}^{\sqrt{y}} f(x)\,dx =$$

Use FTC to get the pdf of Y as

$$g(y) = G'(y) = f(\sqrt{y})\cdot(\sqrt{y})' - f(-\sqrt{y})\cdot(-\sqrt{y})'$$

$$= \frac{1}{\pi(1+(\sqrt{y})^2)}\cdot\frac{1}{2\sqrt{y}} - \frac{1}{\pi(1+(-\sqrt{y})^2)}\cdot\left(-\frac{1}{2\sqrt{y}}\right)$$

$$= \frac{1}{2\sqrt{y}}\,\frac{1}{\pi(1+y)} + \frac{1}{2\sqrt{y}}\,\frac{1}{\pi(1+y)} = \frac{1}{\pi\sqrt{y}\,(1+y)}, \quad y \geq 0$$

Math 3501 - Probability and Statistics I

5.3 - Several random variables

generalizes some of notions from Chp 4 to any number of r.v.s.

Independent random variables

Definition

Let $X_1, X_2, \ldots, X_n$ be n random variables with respective pmfs/pdfs

marginals of $X_1, \ldots, X_n$ $\longrightarrow$ $$f_{X_i}(x_i) \; , \quad x_i \in S_{X_i} \; , \quad i = 1, \ldots, n \; .$$

We say that that $X_1, X_2, \ldots, X_n$ are *independent* if

joint $\longrightarrow$ $$f(x_1, x_2, \ldots x_n) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_n}(x_n)$$

for all $x_1 \in S_{X_1}, x_2 \in S_{X_2}, \ldots, x_n \in S_{X_n}$.

Note: This is the same definition as in the bivariate case:

A sequence of random variables is independent if its joint pmf/pdf equals the product of the respective marginal pmfs/pdfs.

Random sample ← relevant for Statistics (MATH 4501)

Definition

We say that that n random variables $X_1, X_2, \ldots, X_n$ are *identically distributed* if all have the same distribution.

→ $X_1, X_2, \ldots, X_n$ all have the same pmf / pdf

Definition

A collection of n independent and identically distributed (abbreviated i.i.d. or iid) random variables $X_1, X_2, \ldots, X_n$ is said to be a *random sample of size n* from that common distribution.

Note: If $f(x)$ is the common pmf/pdf of these n random variables, then the joint pmf/pdf is

all marginal are the same function

joint pmf

$$f_{\text{joint}}(x_1, x_2, \ldots x_n) = f(x_1)\, f(x_2) \cdots f(x_n) \quad , \quad x_1 \in S_{X_1}, x_2 \in S_{X_2}, \ldots, x_n \in S_{X_n} \ .$$

Interpretation: The values of a random sample may be regarded as measurements on the outcomes of repeated independent random experiments.

Example

Let X_1, X_2, X_3 be a random sample from a distribution with pdf

$$f(x) = e^{-x}, \qquad x > 0 .$$

Determine the joint pdf of X_1, X_2, X_3 and evaluate

$$P(0 < X_1 < 1,\ 2 < X_2 < 4,\ 3 < X_3 < 7) .$$

X_1, X_2, X_3 random sample $\Rightarrow$ X_1, X_2, X_3 are iid with marginal pdf given by f

and so, their joint pdf is

$$f_{joint}(x_1, x_2, x_3) = f(x_1) \cdot f(x_2) \cdot f(x_3) = e^{-x_1} \cdot e^{-x_2} \cdot e^{-x_3}, \quad x_1, x_2, x_3 > 0$$

and so

$$P(0 < X_1 < 1,\ 2 < X_2 < 4,\ 3 < X_3 < 7) = \int_0^1 \int_2^4 \int_3^7 e^{-x_1} \cdot e^{-x_2} \cdot e^{-x_3} \, dx_3 \, dx_2 \, dx_1$$

$$= \left(\int_0^1 e^{-x_1} \, dx_1 \right) \cdot \left(\int_2^4 e^{-x_2} \, dx_2 \right) \cdot \left(\int_3^7 e^{-x_3} \, dx_3 \right)$$

triple integral splits into 3 separate one-dim integrals
due to independence

Mathematical expectation

Definition (Mathematical expectation)

Let [redacted] $X_1, X_2, \ldots, X_n$ be n random variables with joint pmf/pdf $f(x_1, x_2, \ldots, x_n)$ on the space S, and let $u(X_1, X_2, \ldots, X_n)$ be a function of these n random variables.

The *mathematical expectation* (or *expected value*) of $u(X_1, X_2, \ldots, X_n)$ is given by

$$E[u(X_1, X_2, \ldots, X_n)] = \sum_{(x_1, x_2, \ldots, x_n) \in S} u(x_1, x_2, \ldots, x_n) f(x_1, x_2, \ldots, x_n)$$

in the case of discrete random variables, and by

$$E[u(X_1, X_2, \ldots, X_n)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} u(x_1, x_2, \ldots, x_n) f(x_1, x_2, \ldots, x_n) \, \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_n$$

in the case of continuous random variables, both of which provided the multiple sum/integral on the right hand side is absolutely convergent.

Note: This is the same definition as in the bivariate case, with double sums/integrals replaced by the respective multivariate versions.

Expectation of independent random variables

Theorem (identical to the analogue result for the bivariate case)

Let $X_1, X_2, \ldots, X_n$ be independent random variables, and let $u_1, u_2, \ldots, u_n$ be real-valued functions of a single real variable.

The following hold:

1) *The random variables $u_1(X_1), u_2(X_2), \ldots, u_n(X_n)$ are also independent.*
2) *If $E[u_i(X_i)]$ exists for each $i = 1, 2, \ldots, n$, then*

$$E[u_1(X_1)\, u_2(X_2) \cdots u_n(X_n)] = E[u_1(X_1)]\, E[u_2(X_2)] \cdots E[u_n(X_n)]$$

analogue to $E[u(X) \cdot v(Y)] = E[u(X)] \cdot E[v(Y)]$ } Chp 4. Either section 4.1 or 4.2

Linear combination independent random variables

Theorem

Let $X_1, X_2, \dots, X_n$ be n independent random variables with respective means $\mu_1, \mu_2, \dots, \mu_n$ and variances $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$.

The mean and variance of

$$Y = \sum_{i=1}^{n} a_i X_i \ ,$$

where $a_1, a_2, \dots, a_n$ are real constants, are, respectively,

expectation is linear

always holds $\rightarrow$

$$\mu_Y = \sum_{i=1}^{n} a_i \mu_i$$

$\rightarrow E[a_1 X_1 + \dots + a_m X_m] = a_1 E[X_1] + \dots + a_m E[X_m]$

and

when independency is really needed! $\rightarrow$

$$\sigma_Y^2 = \sum_{i=1}^{n} a_i^2 \sigma_i^2 \ .$$

$\rightarrow \mathrm{Var}(a_1 X_1 + \dots + a_m X_m) = a_1^2 \mathrm{Var}(X_1) + \dots + a_m^2 \mathrm{Var}(X_m)$

Example

Let the independent random variables X_1 and X_2 have respective means $\overset{E[X_1]}{\mu_1} = -4$ and $\underset{E[X_2]}{\mu_2} = 3$ and variances $\sigma_1^2 = 4$ and $\sigma_2^2 = 9$.

Find the mean and the variance of $Y = 3X_1 - 2X_2$.

$$E[Y] = E[3X_1 - 2X_2] \underset{\text{linearity}}{=} 3E[X_1] - 2E[X_2] = 3 \cdot \mu_1 - 2 \cdot \mu_2 = 3(-4) - 2(3)$$
$$= \cdots$$

always +

$$\mathrm{Var}(Y) = \mathrm{Var}(3X_1 - 2X_2) \underset{X_1 \text{ and } X_2 \text{ independent}}{=} (3)^2 \mathrm{Var}(X_1) + (-2)^2 \mathrm{Var}(X_2) = 9\sigma_1^2 + 4\sigma_2^2$$
$$= 9(4) + 4(9)$$
$$= 72$$