M378K: February 18th, 2026.

Independence.

Theorem. The Factorization Criterion.

Jointly continuous random variables $Y_1, \dots, Y_n$ are independent

iff

$$f_{Y_1,\dots,Y_n}(y_1,\dots,y_n) = f_{Y_1}(y_1) f_{Y_2}(y_2) \cdots f_{Y_n}(y_n) \text{ for all } y_1,\dots,y_n$$

Corollary. If

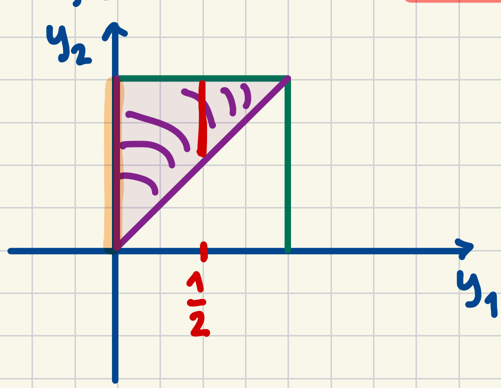
$$f_{Y_1,\dots,Y_n}(y_1,\dots,y_n) = g_1(y_1)\cdot g_2(y_2) \cdots g_n(y_n) \text{ for all } y_1,\dots,y_n$$

and some functions $g_1, \dots, g_n$,

then, r.v.s $Y_1, \dots, Y_n$ are independent.

Example. (Y_1, Y_2) have the joint pdf

$$f_{Y_1,Y_2}(y_1,y_2) = 6y_1 \mathbb{1}_{[0 \le y_1 \le y_2 \le 1]}$$



Theorem. Let $Y_1, \dots, Y_n$ be independent r.v.s and let $h_1, \dots, h_n$ be functions such that

$$h_i(Y_i),\ i = 1..n$$

are all well-defined.

Then, if all the expectations are finite

$$\boxed{\mathbb{E}[h_1(Y_1) \cdot h_2(Y_2) \cdots h_n(Y_n)] = \mathbb{E}[h_1(Y_1)] \cdot \mathbb{E}[h_2(Y_2)] \cdot \ldots \cdot \mathbb{E}[h_n(Y_n)]}$$

e.g., Y_1, Y_n independent

$g_1(y) = g_2(y) = e^y$ for $y \in \mathbb{R}$

$$\mathbb{E}[\exp(Y_1 + Y_2)] = \mathbb{E}[e^{Y_1} \cdot e^{Y_2}] \underset{\uparrow \text{ independence}}{=} \mathbb{E}[e^{Y_1}] \cdot \mathbb{E}[e^{Y_2}]$$

If Y_1 and Y_2 are identically distributed,

i.e., if $F_{Y_1} = F_{Y_2}$,

then,

$$\mathbb{E}[\exp(Y_1 + Y_2)] = (\mathbb{E}[e^{Y_1}])^2$$

M378K Introduction to Mathematical Statistics
Problem Set #8
Transformations of Random Variables.

Problem 8.1. *Let X be a continuous random variable with the cumulative distribution function denoted by F_X and the probability density function denoted by f_X.*

Let the random variable $\boxed{Y = 2X}$ have the p.d.f. denoted by f_Y. Then,

(a) $f_Y(x) = 2f_X(2x)$

(b) $f_Y(x) = \frac{1}{2}f_X\left(\frac{x}{2}\right)$

(c) $f_Y(x) = f_X(2x)$

(d) $f_Y(x) = f_X\left(\frac{x}{2}\right)$

(e) *None of the above*

The CDF Method

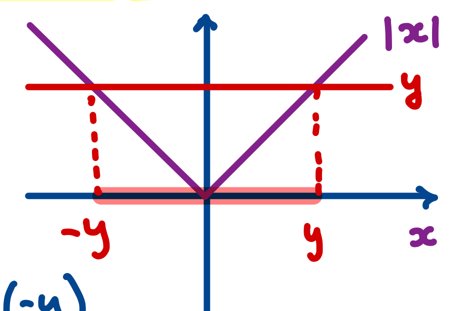
$\longrightarrow$: $y \in \mathbb{R}$: $F_Y(y) = \mathbb{P}[Y \leq y] = \mathbb{P}[2X \leq y]$

$$= \mathbb{P}\left[X \leq \frac{y}{2}\right] = F_X\left(\frac{y}{2}\right)$$

$$f_Y(y) = \frac{d}{dy}F_X\left(\frac{y}{2}\right) = \frac{1}{2}\cdot f_X\left(\frac{y}{2}\right) \quad \square$$

Problem 8.2. *If the $\boxed{\text{continuous}}$ random variable X has the distribution function F_X, then the distribution function of the random variable $Y = |X|$ equals*

$F_Y(y) = ?$



$\longrightarrow$: $F_Y(y) = \mathbb{P}[Y \leq y] = \mathbb{P}[|X| \leq y]$

$$= \mathbb{P}[-y \leq X \leq y] =$$

$$= \mathbb{P}[X \leq y] - \mathbb{P}[X < -y] = F_X(y) - F_X(-y) \quad \square$$

For fun: $f_Y(y) = (f_X(y) + (-1)(-1)f_X(-y))\mathbb{1}_{[0,\infty)}(y) = (f_X(y) + f_X(-y))\mathbb{1}_{[0,\infty)}(y)$

Remark 8.1. *The goal is to figure out the distribution of the random variable*

$$X = g(Y_1, Y_2, \dots, Y_n)$$

where $Y_i, i = 1, \dots, n$ are a **random sample** *with a common density f_Y.*

Def'n. $Y_1, \dots, Y_n$ is a random sample from a dist'n D if:

- $Y_1, \dots, Y_n$ are independent

and

- $Y_i \sim D$ for all $i = 1..n$.

1. *Identify the objective: We want f_X.*
2. *Realize: $f_X = F'_X$*
3. *Recall the definition: $F_X(x) = \mathbb{P}[X \le x] = \mathbb{P}[g(Y_1, \dots, Y_n) \le x]$*
4. *Identify the region A_x in $\mathbb{R}^n$ where*

$$g(y_1, \dots, y_n) \le x$$

   for every x, i.e., express

$$A_x = \{(y_1, \dots, y_n) : g(y_1, \dots, y_n) \le x\}$$

5. *Calculate*

$$F_X(x) = \int \cdots \int_{\mathbb{R}^n} \mathbf{1}_{A_x}(y_1, \dots, y_n) f_Y(y_1) \dots f_Y(y_n) dy_1 \dots dy_n.$$

6. *Differentiate: $f_X = F'_X$.*
7. *Pat yourself on the back!*

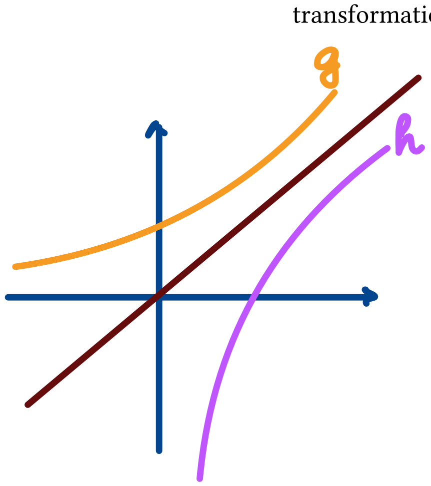
Problem 8.3. **One-to-one transformations: Step-by-step** *Let Y be a random variable with density f_Y. Let $g : \mathbb{R} \to \mathbb{R}$ be a strictly increasing differentiable function. Define $\tilde{Y} = g(Y)$. What is the density function $f_{\tilde{Y}}$ of $\tilde{Y}$ expressed in terms of f_Y and g?*

1. Identify the objective: We want $f_{\tilde{Y}}$.
2. Realize: $f_{\tilde{Y}} = F'_{\tilde{Y}}$
3. Recall the definition:

$$F_{\tilde{Y}}(x) = \mathbb{P}[\tilde{Y} \le x] = \mathbb{P}[g(Y) \le x]$$

4. The function g is assumed to be **strictly increasing**. In which way can you modify the inequality in the probability you obtained above to *separate* the random variable Y from the transformation g?

There exists $h = g^{-1}$.
This is g's INVERSE FUNCTION.
h is also increasing



$$F_{\tilde{Y}}(x) = \mathbb{P}[g(Y) \le x] = \mathbb{P}[\cancel{h}(\cancel{g}(Y)) \le h(x)]$$
$$= \mathbb{P}[Y \le h(x)]$$

5. Express your result from above in terms of the c.d.f. F_Y of the r.v. Y.

$$F_{\tilde{Y}}(x) = \mathbb{P}[Y \leq h(x)] = F_Y(h(x))$$

6. Differentiate: $f_{\tilde{Y}} = F'_{\tilde{Y}}$.

$$f_{\tilde{Y}}(x) = \frac{d}{dx} F_{\tilde{Y}}(x) = \frac{d}{dx} F_Y(h(x)) = \text{(chain rule)}$$
$$= f_Y(h(x)) \cdot h'(x) \qquad \square$$

Problem 8.4. *The time T that a manufacturing distribution system is out of operation is modeled by a distribution with the following c.d.f.*

$$F_T(t) = (1 - (2/t)^2)\mathbf{1}_{(2,\infty)}(t)$$

The resulting cost to the company is $Y = T^2$. Find the probability density function f_Y of the r.v. Y.

Problem 8.5. *What if h is strictly decreasing?*

Problem 8.6. *The unifying formula?*