M378K: February 14th, 2025.

Independence [cont'd].

Let $Y_1, \dots, Y_n$ be independent r.v.s.

Let $g_1, \dots, g_n$ be functions such that $g_i(Y_i)$, $i=1 \dots n$, are all well defined.

Then, if all expectations are finite,

$$\boxed{\mathbb{E}[g_1(Y_1) \cdot g_2(Y_2) \cdots g_n(Y_n)] = \mathbb{E}[g_1(Y_1)] \cdot \mathbb{E}[g_2(Y_2)] \cdots \mathbb{E}[g_n(Y_n)]}$$

e.g., Y_1, Y_2 independent

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

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

If Y_1 and Y_2 are also identically distributed,

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

Def'n. Y_1 and Y_2 are said to be identically distributed if

$$\boxed{F_{Y_1} = F_{Y_2}}$$

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

$\boxed{\text{The CDF Method}}$

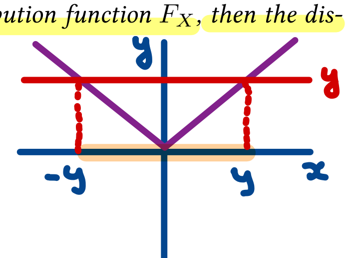
→: $y \in \mathbb{R}$: $F_Y(y) = \mathbb{P}[Y \leq y] = \mathbb{P}[2X \leq y] =$

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

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

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

$F_Y(y) = ?$



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

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

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

$f_Y(y) = \left(f_X(y) + f_X(-y)\right) \mathbb{1}_{[0,\infty)}(y)$ □

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 distribution D if:
(i) $Y_1, \dots, Y_n$ are independent,
(ii) $Y_i \sim D$ for all $i = 1..n$

1. *Identify the objective: We want f_X.*

2. *Realize: $f_X = F_X'$*

3. *Recall the definition: $F_X(x) = \mathbb{P}[X \le x] = \mathbb{P}[g(Y_1, \ldots, Y_n) \le x]$*

4. *Identify the region A_x in $\mathbb{R}^n$ where*

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

   for every x, i.e., express

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

5. *Calculate*

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

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 $\boxed{\tilde{Y} = g(Y)}$. What is the density function $f_{\tilde{Y}}$ of $\tilde{Y}$ expressed in terms of f_Y and g?*

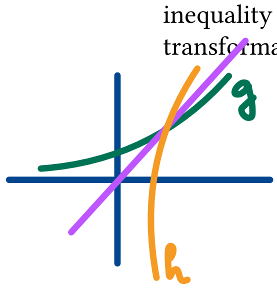
1. Identify the objective: We want $f_{\tilde{Y}}$.

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?



exists $h = g^{-1}$

This is g's INVERSE FUNCTION; it's also strictly increasing.

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

$$= \mathbb{P}[Y \le \boxed{h(x)}]$$

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

$$F_{\tilde{Y}}(x) = F_Y(h(x))$$

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

$$f_{\tilde{Y}}(x) = \frac{d}{dx} F_Y(h(x)) = f_Y(h(x))\, h'(x)$$

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.*

$$F_T(t) = (1 - (2/t)^2)\mathbf{1}_{(2,\infty)}(t) = \begin{cases} 1 - 4t^{-2} & t > 2 \\ 0 & t \leq 2 \end{cases}$$

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

$\longrightarrow$: $g(t) = t^2,\ t > 2 \Rightarrow \boxed{h(y) = \sqrt{y}},\ y > 4 \Rightarrow h'(y) = \frac{1}{2\sqrt{y}},\ y > 4$

$f_T(t) = \frac{8}{t^3} \mathbb{1}_{(2,\infty)}(t)$

$$f_Y(y) = \frac{\cancel{8}^{\,4}}{(\sqrt{y})^3} \cdot \frac{1}{\cancel{2}\sqrt{y}} \mathbb{1}_{(4,\infty)}(y) = \frac{4}{y^2} \mathbb{1}_{(4,\infty)}(y) \quad \square$$

Problem 8.5. *What if h is $\boxed{\text{strictly decreasing}}$?*

$\boxed{h = g^{-1}}$

$$F_{\tilde{Y}}(y) = \mathbb{P}[\tilde{Y} \leq y] = \mathbb{P}[g(Y) \leq y] \,!\!= \mathbb{P}[Y \geq h(y)] =$$
$$= 1 - \mathbb{P}[Y \leq h(y)] = 1 - F_Y(h(y))$$

$x_1 < x_2 \Rightarrow g(x_1) > g(x_2)$

$$f_{\tilde{Y}}(y) = -f_Y(h(y)) \cdot \underbrace{h'(y)}_{<0}$$

Problem 8.6. *The unifying formula?*

$$\boxed{f_{\tilde{Y}}(y) = f_Y(h(y)) \cdot |h'(y)|}$$