

## M378K Introduction to Mathematical Statistics

### Problem Set #6

#### Transformations of Random Variables.

---

**Problem 6.1.** 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  $Y = 2X$  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

**Solution: (b)**

For every  $x \in \mathbb{R}$ , the cumulative distribution function is

$$F_Y(x) = \mathbb{P}[Y \leq x] = \mathbb{P}[2X \leq x] = \mathbb{P}[X \leq x/2] = F_X(x/2).$$

As for the probability density function, we have that for all  $x$ ,

$$f_Y(x) = F'_Y(x) = f_X(x/2)/2.$$

**Problem 6.2.** If the continuous random variable  $X$  has the distribution function  $F_X$ , then the distribution function of the random variable  $Y = |X|$  equals

$$F_Y(y) = ?$$

**Solution:**

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

**Remark 6.1.** 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  $f_Y$ .

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

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

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

for every  $x$ , i.e., express

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

5. Calculate

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

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

**Problem 6.3. One-to-one transformations: Step-by-step** Let  $Y$  be a random variable with density  $f_Y$ . Let  $g : \mathbb{R} \rightarrow \mathbb{R}$  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:

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

**Solution:**

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

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  $g$ ?

**Solution:** If  $g$  is strictly increasing, then it is one-to-one and  $h = g^{-1}$  exists (and it is also **increasing**). So,

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

Note that the direction of the inequality remains unchanged.

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

**Solution:**

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

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

**Solution:** The inverse  $h^{-1}$  is differentiable since  $h$  is differentiable. So,

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

**Problem 6.4.** 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.  $Y$ .

**Solution:** In this problem

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

and

$$g(x) = x^2, x > 2 \quad \Rightarrow \quad h(x) = \sqrt{x}, \quad x > 4.$$

The derivative:

$$\frac{d}{dx} h(x) = \frac{1}{2\sqrt{x}} \quad x > 4.$$

Finally,

$$f_Y(y) = \frac{1}{2\sqrt{y}} \times \frac{8}{(\sqrt{y})^3} = \frac{4}{y^2} \quad y > 4.$$

**Problem 6.5.** What if  $h$  is strictly decreasing?

**Solution:**

$$F_{\tilde{Y}}(y) = 1 - F_Y(h(y))$$

So,

$$f_{\tilde{Y}}(y) = -\frac{d}{dy} h(y) f_Y(h(y))$$

**Problem 6.6.** The unifying formula?

**Solution:**

$$f_Y(y) = \left| \frac{d}{dy} h(y) \right| f_X(y)$$

Do not forget: it always makes sense to simply attack a problem without giving it a “label” .... Just look at the following problem:

**Problem 6.7.** Let  $T_1$  and  $T_2$  be independent shifted geometric random variables with parameters  $p_1 = 1/2$  and  $p_2 = 1/3$ . Compute  $\mathbb{E}[\min(T_1, T_2)]$ .

**Solution:** By the assumption, for  $k \geq 1$ , we have

$$\mathbb{P}[T_1 \geq k] = \sum_{i=k}^{\infty} (1-p_1)^{i-1} p_1 = (1-p_1)^{k-1} = \frac{1}{2^{k-1}}.$$

Similarly, we have  $\mathbb{P}[T_2 \geq k] = (1-p_2)^{k-1} = (2/3)^{k-1}$ , and so

$$\begin{aligned} \mathbb{P}[\min(T_1, T_2) \geq k] &= \mathbb{P}[T_1 \geq k \text{ and } T_2 \geq k] \\ &= \mathbb{P}[T_1 \geq k] \times \mathbb{P}[T_2 \geq k] \\ &= \frac{1}{3^{k-1}} = \left(1 - \frac{2}{3}\right)^{k-1}. \end{aligned}$$

It follows that  $\min(T_1, T_2)$  is geometrically distributed with  $p = 2/3$ , and, so  $\mathbb{E}[\min(T_1, T_2)] = 3/2$ .