Lectures 11-12

A Classification of random variables. Random variables can be classified using distribution functions according to their 'continuity properties.

Definition 5.4: A random variable X with distribution function $F : \mathbb{R} \to \mathbb{R}$ is said to be a discrete random variable if

$$\sum_{x \in D} \big(F(x) \; - \; F(x-) \big) \; = \; 1 \, ,$$

where D is the set of discontinuities of F.

Here observe that the 'discrete' distribution F 'exhaust' all the probability masses through the jumps.

Lemma 0.1 *If F is the distribution of a discrete ranodm variable X, then it is of the form*

$$F(x) \;\; = \sum_{i : x_i \in D} p_i H_0(x - x_i), \;\; x \in \mathbb{R},$$

where H_0 denote the Heaviside function[1] *and $p_i = P\big(X = x_i\big) = F(x_i) - F(x_i-)$ and D is the set of discontinuities of F.*

Proof: Let $D = \{x_i | i \in I\}$ where the index set I is countable and X be a random variable with distribution F. Then it follows that $P\big(X = x\big) = 0$ for all $x \notin D$. For $x \in \mathbb{R}$, let $I_x \;=\; \{i \in I | x_i \leq x\}$. Then observe that $i \in I_x \Leftrightarrow H_0(x - x_i) = 1$ and $i \notin I_x \Leftrightarrow H_0(x - x_i) = 0$. Hence

$$\begin{aligned}
F(x) &= P\big(X \leq x\big) \\
&= \sum_{i \in I_x} P\big(X = x_i\big) \;=\; \sum_{i \in I_x} p_i \times 1 + \sum_{i \notin I_x} p_i \times 0 \\
&= \sum_{i \in I} p_i H_0(x - x_i).
\end{aligned}$$

The distributions in Examples 4 to 8 from Lecture notes 9-10 corresponds to discrete random variables.

[1] Heaviside function H_0 is defined by

$$H_0(x) \;=\; \begin{cases} 0 & \text{if} \quad x < 0 \\ 1 & \text{if} \quad x \geq 0. \end{cases}$$

Definition 5.5 A random variable X with distribution function $F : \mathbb{R} \to \mathbb{R}$ which is continuous is said to be random variable with continuous distribution and is in short called by the name continuous random variable.

The distributions given in Examples 9 to 11 from Lecture notes 9-10 corresponds to continuous random variables.

Definition 5.6 (Probability mass function)

Let X be a discrete random variable with distribution function $F : \mathbb{R} \to \mathbb{R}$.

Define $f : \mathbb{R} \to \mathbb{R}$ as follows:

$$f(x) \ = \ F(x) \ - \ F(x-)$$

Then f is called the probability mass function(pmf) of X.

For example, the pmf of the discrete random variable of Example **??** is given by

$$f(x) \ = \ \begin{cases} \frac{1}{4} & \text{if} \quad x = 0, 2 \\ \frac{1}{2} & \text{if} \quad x = 1 \\ 0 & \quad \text{otherwise} \,. \end{cases}$$

It is left as an exercise for the student to write down the pmf of random variables in Examples 4 to 8 from Lecturenotes 9-10.

The pmf of a continuous random variable is the zero function. Hence the notion of pmf is useless for continuous random variables.

Definition 5.7(Probability density function)

A continuous random variable X with distribution function $F : \mathbb{R} \to \mathbb{R}$ is said to have a probability density function(pdf) if there exists a function $f : \mathbb{R} \to \mathbb{R}$ such that

$$F(x) \ = \ \int_{-\infty}^{x} f(y)dy \ \ \forall \ \ x \in \mathbb{R}$$

If $f : \mathbb{R} \to \mathbb{R}$ exists, then it is called the pdf of X.

A continuous random variable with a pdf is called **absolutely continuous** random variable

It is easy to see that if F is differentiable every where and the derivative denoted by F' is a continuous function, then the corresponding random variable X has a pdf and is given by $f = F'$. This is not a necessary condition.

Example 0.1 *Define $F : \mathbb{R} \to \mathbb{R}$ as follows.*

$$F(x) = \begin{cases} 0 & \text{if } \ x < 0 \\ x & \text{if } \ 0 \leq x < \frac{1}{2} \\ \frac{1}{2} & \text{if } \ \frac{1}{2} \leq x < 1 \\ x - \frac{1}{2} & \text{if } \ 1 \leq x < \frac{3}{2} \\ 1 & \text{if } \ x \geq \frac{3}{2}\,. \end{cases}$$

Student can verify that F is a distribution function as it satisfies the conditions (1) and (2) given in the definition 5.3. Also verify that it corresponds to the random variable given by the random experiment of picking a point 'at random' from $[0,\ \frac{1}{2}] \cup [1,\ \frac{3}{2}]$.

F is not differentiable at $x = \frac{1}{2}, 1$. Verify that the function

$$f(x) = \begin{cases} 0 & \text{if } \ x < 0 \\ 1 & \text{if } \ 0 \leq x < \frac{1}{2} \\ 0 & \text{if } \ \frac{1}{2} \leq x < 1 \\ 1 & \text{if } \ 1 \leq x < \frac{3}{2} \\ 0 & \text{if } \ x \geq \frac{3}{2}\,. \end{cases}$$

is the pdf of F (exercise).

Consider the distribution function

$$\begin{aligned} F(x) &= 0 \quad \text{if } \ x < 0 \\ &= \tfrac{1}{4} \quad \text{if } \ 0 \leq x < 1 \\ &= \tfrac{3}{4} \quad \text{if } \ 1 \leq x < 2 \\ &= 1 \quad \text{if } \ x \geq 2\,. \end{aligned}$$

Note that F is differentiable except at the points $0, 1$, and 2 and $F'(x) = 0$ for $x \neq 0, 1, 2$. Here the pdf doesn't exists. Compare this with the distribution function in Example 0.1.

Distribution function of transformation of random variables: In this subsection, we will see how one can write down the distribution function of $Y = \varphi \circ X$ in terms of the distribution of X where $\varphi : \mathbb{R} \to \mathbb{R}$ is Borel measurable. Note it is not possible to give an explicit formula but in some cases one will be able to do that. Here my plan is to give a general recipe and will illustrate it through some examples. I will give as an example, one special class of transformations. Though it is possible to give an explicit formula for many other cases, I will not do it instead take some examples and show you how to use it.

We use μ_Z, F_Z to denote the distribution and distribution function of a random variable Z from now onwards.

General Recipe: The distribution of Y is given by

$$\begin{aligned}\mu_Y(B) &= P\big(Y \in B\big)\\ &= P\big(\varphi(X) \in B\big)\\ &= P\big(X \in \varphi^{-1}(B)\big)\\ &= \mu_X\big(\varphi^{-1}(B)\big),\ B \in \mathcal{B}_{\mathbb{R}}.\end{aligned}$$

Hence by taking $B = (-\infty, y],\ y \in \mathbb{R}$, we get following:

$$F_Y(y) \ = \ \mu_X\big(\varphi^{-1}(-\infty,\ y]\big),\ y \in \mathbb{R}.$$

Hence to compute the distribution function Y in terms of the distribution function of X, one need to identify the set $\varphi^{-1}(-\infty, y])$. This I will illustrate in the next example.

Lemma 0.2 *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a continuous function which is non decreasing. Then $\varphi^{-1}(-\infty, y]) \ = \ (-\infty,\ \sup \varphi^{-1}(y)]$. This implies that*

$$F_Y(y) \ = \ F_X(\sup \varphi^{-1}(y)),\ y \in \mathbb{R}.$$

In particular, if φ is strictly increasing, then $F_Y(y) \ = \ F_X(\varphi^{-1}(y))$.

Proof: Now we will see the proof of $\varphi^{-1}(-\infty, y]) \ = \ (-\infty,\ \sup \varphi^{-1}(y)]$.

$$\begin{aligned}x \in \varphi^{-1}(-\infty, y]) &\Rightarrow \varphi(x) \in (-\infty, y]\\ &\Rightarrow \varphi(x) < y \text{ or } \varphi(x) = y\\ &\Rightarrow x < z \text{ for all } z \in \varphi^{-1}(y) \text{ or } \varphi(x) = y\\ &\Rightarrow x \in (-\infty,\ \sup \varphi^{-1}(y)].\end{aligned}$$

The proof of statement $\varphi(x) < y \Rightarrow x < z$ for all $z \in \varphi^{-1}(y)$ is as follows. Suppose there exists some $z \in \varphi^{-1}(y)$ such that $x \geq z$, then $\varphi(x) \geq \varphi(z) = y$, a contradiction to $\varphi(x) < y$.

Now we prove the reverse inclusion. Suppose $x \leq \sup \varphi^{-1}(y)$. Then either (I) : $x < \sup \varphi^{-1}(y) \Rightarrow x \leq z$ for some $z \in \varphi^{-1}(y)$
or (II) : $x = \sup \varphi^{-1}(y) \Rightarrow x \geq z$ for all $z \in \varphi^{-1}(y)$ and there exists a sequence $z_n \in \varphi^{-1}(y)$ satisfying $z_n \to x$.

Now

$$\begin{aligned} \text{(I)} \ &\Rightarrow \ \varphi(x) \leq y \\ &\Rightarrow \ x \in \varphi^{-1}((-\infty, y]). \end{aligned}$$

$$\begin{aligned} \text{(II)} \ &\Rightarrow \ \varphi(x) = \lim_{n\to\infty} \varphi(z_n) = y \text{ (using continuity of } \varphi) \\ &\Rightarrow \ \varphi(x) = y \\ &\Rightarrow \ x \in \varphi^{-1}(y) \subseteq \varphi^{-1}((-\infty, y]). \end{aligned}$$

This completes the proof of the reverse inclusion. Hence the proof is complete. □

Lemma 0.3 *Let X be a random variable with pdf f such that $f \equiv 0$ outside and interval I and $\varphi : I \to \mathbb{R}$ is a strictly monotone and differentiable function. Then pdf g of $\varphi \circ X$ exists and is given by*

$$g(y) \ = \ f(\varphi^{-1}(y))|(\varphi^{-1})'(y)|, y \in \varphi(I), \ g(y) = 0 \text{ elsewhere.}$$

Proof: When φ is strictly increasing, and let a, b denote respectively the left and right endpoints of I. For $y \in \varphi(I)$,

$$\begin{aligned} F_Y(y) &= P\big(\varphi \circ X \leq y\big) \\ &= P\big(X \leq \varphi^{-1}(y)\big) \\ &= F_X(\varphi^{-1}(y)). \end{aligned}$$

Hence

$$F_Y(y) \ = \ \begin{cases} 0 & \text{if} \ \ y < \varphi(a) \\ F_X(\varphi^{-1}(y)) & \text{if} \ \ \varphi(a) \leq y < \varphi(b) \\ 1 & \text{if} \ \ y \geq \varphi(b). \end{cases}$$

Now for $\varphi(a) \leq y < \varphi(b)$,

$$\begin{aligned} F_Y(y) &= F_X(\varphi^{-1}(y)) \\ &= \int_{-\infty}^{\varphi^{-1}(y)} f(t)dt = \int_a^{\varphi^{-1}(y)} f(t)dt \\ (\text{put } t = \varphi^{-1}(u)) &= \int_{\varphi(a)}^{y} f(\varphi^{-1}(u))\, (\varphi^{-1})'(u)du \\ &= \int_{-\infty}^{y} g(u)du, \end{aligned}$$

where

$$g(y) \;=\; \begin{cases} f(\varphi^{-1}(y))\,(\varphi^{-1})'(y) & \text{if} \qquad \varphi(a) \leq y < \varphi(b) \\ 0 & \text{otherwise} \end{cases}$$

For $y \notin [\varphi(a),\ \varphi(b))$, it is easy to see that

$$F_Y(y) = \int_{-\infty}^{y} g(u)du.$$

This completes the proof. □

Example 0.2 $\varphi(x) = x^3$. *(Prototype for φ which is increasing and continuous) Hence $F_Y(y) = F_X(y^{\frac{1}{3}})$. Here $y^{\frac{1}{3}}$ denote the real cube root of y.*

Example 0.3 *Let X be Uniform $(-\frac{\pi}{2},\ \frac{\pi}{2})$ random variable and $Y = \tan X$.*

Here note $\varphi(x) \;=\; \tan x,\ x \in (-\frac{\pi}{2},\ \frac{\pi}{2})$ is strictly increasing, differentiable and $\varphi(I) = \mathbb{R}$. Hence using Lemma 0.3, it follows that Y has a pdf g given by

$$\begin{aligned} g(y) &= f(\tan^{-1} y)\frac{d}{dy}\tan^{-1} y,\ y \in \mathbb{R} \\ &= \frac{1}{\pi}\frac{1}{1+y^2},\ y \in \mathbb{R}, \end{aligned}$$

where $f(x) = \frac{1}{\pi}, x \in (-\frac{\pi}{2},\ \frac{\pi}{2})$ and $= 0$ otherwise. Hence Y is a Cauchy random variable.

Example 0.4 *Let $\varphi(x) = x^2 + 1$. (Prototype for convex functions). Then*

$$\varphi^{-1}(-\infty, y] \;=\; \begin{cases} \emptyset & \text{if} \quad y < 1 \\ \{0\} & \text{if} \quad y = 1 \\ [-\sqrt{y-1},\ \sqrt{y-1}] & \text{if} \quad y > 1. \end{cases}$$

Hence

$$F_Y(y) \;=\; \mu_X([-\sqrt{y-1},\ \sqrt{y-1}]) \;=\; F_X(\sqrt{y-1}) - F_X(\sqrt{y-1}-).$$

Warning : Don't use Lemma 5.3 in this case as φ is not monotone.

Example 0.5 *φ be the Heaviside function, i.e. $\varphi(x) = 0$ if $x < 0$ and $= 1$ if $x \geq 0$. (Prototype for φ which is piece-wise continuous)Then*

$$\varphi^{-1}(-\infty, y]) \;=\; \begin{cases} \emptyset & \text{if} \quad y < 0 \\ (-\infty,\ 0) & \text{if} \quad 0 \leq y < 1 \\ \mathbb{R} & \text{if} \quad y \geq 1. \end{cases}$$

Hence

$$F_Y(y) = \begin{cases} 0 & \text{if } \ y < 0 \\ F_X(0-) & \text{if } \ 0 \le y < 1 \\ 1 & \text{if } \ y \ge 1. \end{cases}$$

Example 0.6 *Let X be a normal with parameters 0 and σ and $Y = X^2$, i.e. $\varphi(x) = x^2, Y = \varphi \circ X$. Then*

$$\varphi^{-1}(-\infty, y] = \begin{cases} \emptyset & \text{if } \ y < 0 \\ [-\sqrt{y}, \ \sqrt{y}] & \text{if } \ y \ge 0. \end{cases}$$

Hence

$$F_Y(y) = F_X(\sqrt{y}) - F_X(-\sqrt{y}), y \ge 0.$$

Hence the pdf g of Y is given by

$$g(y) = \frac{1}{2\sqrt{y}}(f(\sqrt{y}) + f(-\sqrt{y}), y \ge 0,$$

where f is the pdf of X and is given by

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}}, x \in \mathbb{R}.$$

Hence

$$g(y) = \frac{1}{\sqrt{2\pi y\sigma}} e^{-\frac{y}{2\sigma^2}}, y > 0,$$

i.e., Y follows Gamma distribution with parameters $\frac{1}{2}$ and $\frac{1}{2\sigma^2}$.

Recall that a random variable X follows gamma with parameters (α, λ) if its pdf is given by

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, x > 0, \ g(y) = 0, \ x \le 0,$$

where

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx, \alpha > 0.$$

Example 0.7 *Let X be uniform $(0, \ 1)$ and $Y = \max\{\frac{1}{2}, X\}$. Then $\varphi(x) = \max\{\frac{1}{2}, x\}, x \in \mathbb{R}$ an increasing function. Here*

$$\varphi^{-1}(y) = \begin{cases} \emptyset & \text{if } \ y < \frac{1}{2} \\ (-\infty, \ \frac{1}{2}] & \text{if } \ y = \frac{1}{2} \\ \{y\} & \text{if } \ y > \frac{1}{2}. \end{cases}$$

$$F_Y(y) \;=\; F_X(\sup \varphi^{-1}(y)) = \begin{cases} 0 & \text{if} \quad y < \frac{1}{2} \\ F_X(\frac{1}{2}) & \text{if} \quad y = \frac{1}{2} \\ F_X(y) & \text{if} \quad y > \frac{1}{2}. \end{cases}$$

Hence

$$F_y(y) \;=\; \begin{cases} 0 & \text{if} \quad y < \frac{1}{2} \\ \frac{1}{2} & \text{if} \quad y = \frac{1}{2} \\ y & \text{if} \quad \frac{1}{2} < y < 1 \\ 1 & \text{if} \quad y \geq 1. \end{cases}$$

Chapter 5: Random vectors, joint distributions

Key words: Random vector, Borel σ-filed of subsets of $\mathbb{R}^n$, joint distribution function, law of random vector or joint distribution, marginal distribution functions, conditional pmf, conditional pdf.

In many problems, one often encounter multiple random objects. For example, future price of two different stocks in a stock market. Since the price of one stock can affect the price of the second, it is not advisable to analyze them separately. One should to analyze multiple random objects together, since behavior of one can affect others. To model such phenomenon, we need to introduce many random variables in a single platform (i.e., a probability space).

First we will recall, some elementary facts about n-dimensional Euclidean space. Consider

$$\mathbb{R}^n \;=\; \overbrace{\mathbb{R} \times \cdots \times \mathbb{R}}^{n\, times}$$

with the usual metric $d(x, y) \;=\; \|x - y\|$, where

$$\|x\| \;=\; \sqrt{\sum_{i=1}^{n} |x_i|^2}, \;\; x = (x_1, \ldots . x_n) \in \mathbb{R}^n \, .$$

A subset O of $\mathbb{R}^n$ is said to be open if for each $x \in O$, there exists an $\epsilon > 0$ such that

$$B(x, \epsilon) \subseteq O,$$

where

$$B(x, \epsilon) \;=\; \{y \in \mathbb{R}^n \,|\, \|x - y\| < \epsilon\} \, .$$

i.e., each point in an open set is an interior point. A useful fact: Any open set can be written as a countable union of open sets of the form $\Pi_{i=1}^n (a_i, \; b_i), \; a_i < b_i$, they are called open rectangles.

Definition 6.1. The σ-field generated by all open sets in $\mathbb{R}^n$ is called the Borel σ-field of subsets of $\mathbb{R}^n$ and is denoted by $\mathcal{B}_{\mathbb{R}^n}$.

Some Borel sets : Rectangles, triangles, lines, points all are Borel sets in $\mathbb{R}^2$.

Theorem 0.1 *Let*

$$\mathcal{I}_n \;=\; \{(-\infty,\ x_1] \times \cdots \times (-\infty,\ x_n] \,|\, (x_1, \ldots, x_n) \in \mathbb{R}^n\}\,.$$

Then

$$\sigma(\mathcal{I}_n) \;=\; \mathcal{B}_{\mathbb{R}^n}\,.$$

Proof. (Reading exercise) We prove for $n = 2$, for $n \geq 3$, it is similar. Note that

$$\mathcal{I}_2 \;\subseteq\; \mathcal{B}_{\mathbb{R}^2}\,.$$

Hence from the definition of $\sigma(\mathcal{I}_2)$, we have

$$\sigma(\mathcal{I}_2) \;\subseteq\; \mathcal{B}_{\mathbb{R}^2}\,.$$

Note that for $(x_1, x_2) \in \mathbb{R}^2$,

$$(-\infty,\ x_1) \times (-\infty,\ x_2) \;=\; \bigcup_{m=1}^{\infty} \left[\left(-\infty,\ x_1 - \frac{1}{m}\right] \times \left(-\infty,\ x_2 - \frac{1}{m}\right]\right] \in \sigma(\mathcal{I}_2)\,.$$

For each $x_1, x_2, y_1, y_2 \in \mathbb{R}$ such that $x_1 < y_1, x_2 < y_2$ we have

$$\begin{aligned}(x_1,\ y_1) \times (x_2,\ y_2) \;&=\; (-\infty,\ y_1) \times (-\infty,\ y_2) \\ &\quad \setminus \Big[(-\infty,\ x_1] \times (-\infty,\ x_2] \cup (-\infty,\ x_1] \times (-\infty,\ y_2] \\ &\quad \cup (-\infty,\ y_1] \times (-\infty,\ x_2]\Big]\,.\end{aligned}$$

Hence all bounded open rectangles are in $\sigma(\mathcal{I}_2)$. Since any open set in $\mathbb{R}^2$ can be rewritten as a countable union of bounded open rectangles, all open sets are in $\sigma(\mathcal{I}_2)$. Therefore from the definition of $\mathcal{B}_{\mathbb{R}^2}$, we get

$$\mathcal{B}_{\mathbb{R}^2} \;\subseteq\; \sigma(\mathcal{I}_2)\,.$$

This completes the proof. (It is advised that student try to write down the proof for $n = 3$)