Distributions of Random Variables

(ref: Rosenthal §6, Billingsley §20)

<u>def</u>: Given a random variable X defined on a probability space $(\Omega, \mathcal{F}, P)$, its <u>distribution</u> (or <u>law</u>) is the function μ defined on $\mathcal{B}$ (Borel subsets of $\mathbb{R}$) by

$$\mu(B) = P(X \in B) = P(X^{-1}(B)) \quad \text{for } B \in \mathcal{B}.$$

<u>Notation</u>: If μ is the distribution of X, then

- $(\mathbb{R}, \mathcal{B}, \mu)$ is a valid probability space
- Sometimes write μ as $\mathcal{L}(X)$ for "law" of X
- Write $X \sim \mu$ for "μ is the distribution of X" or "X follows distribution μ"

<u>def</u>: The <u>cumulative distribution function</u> of a RV X by $\boxed{F_X(x) = P(X \le x)}$ for $x \in \mathbb{R}$.

Properties of CDF

- By continuity of probabilities, F_X is right-continuous:
  i.e. if $\{x_n\} \searrow x$ then $F_X(x_n) \to F_X(x)$.

- F_X is a non-decreasing function of x, with

$$\lim_{x\to\infty} F_X(x) = 1 \quad \text{and} \quad \lim_{x\to-\infty} F_X(x) = 0$$

Prop: Let X & Y be two RVs (possibly defined on different prob. spaces). Then

$$\mathcal{L}(X) = \mathcal{L}(Y) \quad \text{iff} \quad F_X(x) = F_Y(x) \text{ for } x \in \mathbb{R}.$$

The following theorem shows that distributions completely specify the expected values of RVs (& functions of them).

Thm [Change of Variable Theorem]:

Given a probability space $(\Omega, \mathcal{F}, P)$, let X be a RV having distribution μ. Then for any Borel-measurable function $f: \mathbb{R} \to \mathbb{R}$, we have

Notation: or $P(d\omega)$ ... $\mu(dt)$

(6.1.2) $$\boxed{\int_\Omega f(X(\omega))\, dP(\omega) = \int_{-\infty}^{\infty} f(t)\, d\mu(t)}$$

i.e. $E_P[f(X)] = E_\mu[f]$ provided that either side is well-defined.

Properties of CDF - extra details

- By continuity from above, F is right-continuous.

$$\{x_n\} \searrow x \implies F_X(x_n) \to F_X(x) \text{ as } n \to \infty$$

- Since F is non-decreasing function of x, the left-hand limit

$$F_X(x-) = \lim_{y \to x^-} F_X(x_n) \implies F_X(x-) = P(X < x)$$

i.e. $\{x_n\} \nearrow x$

So by continuity from below for μ,
the jump in F_X at x is

$$F_X(x) - F_X(x-) = \mu(\{x\}) = P(X = x).$$

* Therefore, F can have at most countably many points of discontinuity (by Thm 10.2 (iv) in B).

since μ is a σ-finite measure, i.e.

$\mu(\Omega) = 1 < \infty$

probability measure

finite & σ-finite measure

In words,

the exp. value of RV $f(X)$ w.r.t. the prob. measure P on Ω is equal to the exp. value of the function f w.r.t. the measure μ on $\mathbb{R}$.

Cor: Let X & Y be 2 RVs (possibly defined on different prob. spaces). Then $\mathcal{L}(X) = \mathcal{L}(Y)$ iff

$$E[f(X)] = E[f(Y)] \quad \forall \text{ Borel-measurable functions}$$

$f: \mathbb{R} \to \mathbb{R}$ for which either expectation is well-defined.

Cor: If X & Y are RVs s.t. $P(X=Y)=1$, then

$$E[f(X)] = E[f(Y)] \quad \forall \text{ Borel-meas. functions } f: \mathbb{R} \to \mathbb{R}$$

for which either expectation is well-defined.

(If $\mu = \mathcal{L}(X) = \mathcal{L}(Y)$, then $E[f(X)] = E[f(Y)] = \int_{\mathbb{R}} f\, d\mu$)

Pf [Change of Var. Thm]: First suppose that $f = \mathbb{1}_B$ for $\underset{\text{Borel set on } \mathbb{R}}{B} \in \mathcal{B}$.

Then $\int_\Omega f(X(\omega))\, dP(\omega) = \int_\Omega \mathbb{1}_{\{X(\omega) \in B\}}\, dP(\omega) = P(X \in B)$,

while $\int_{-\infty}^{\infty} f(t)\, d\mu(t) = \int_{-\infty}^{\infty} \mathbb{1}_{\{t \in B\}}\, d\mu(t) = \mu(B) = P(X \in B)$.

Hence, equality holds in this case.

Now suppose that f is a non-neg. simple function.

Then f is a finite positive linear combination of indicator functions. Both sides of (6.1.2) are linear functions of f, so equality holds in this case.

Next suppose that f is a general non-neg. Borel measurable function. Then we can find a sequence $\{f_n\}$ of non-neg. simple functions s.t. $\{f_n\} \nearrow f$. We know (by above argument) that (6.1.2) holds when f is replaced by f_n. Let $n \to \infty$ & then MCT $\Rightarrow$ (6.1.2) holds for f as well.

Finally, for general Borel-meas. f, write $f = f^+ - f^-$. Since (6.1.2) holds for f^+ & f^- separately & since f is linear, it must also hold for f. ∎

* This proof method is widely used!

- indicator functions
- non-neg. simple fct
- non-neg. general fct
- general functions

Examples of Distributions

Ex 1: RV X s.t. $P(X = \underset{\text{constant}}{c}) = 1$ for some $c \in \mathbb{R}$

The distribution of X is the point mass δ_c defined by $\delta_c(B) = \mathbb{1}_B(c)$

i.e. $\delta_c(B) = \begin{cases} 1 & \text{if } c \in B \\ 0 & \text{if } c \notin B \end{cases}$

Write $X \sim \delta_c$ or $\boxed{\mathcal{L}(X) = \delta_c}$

$E[X] = E[c] = c$

In general, $E[f(X)] = f(c)$ for any function f.

$$\underbrace{\int_\Omega f(X(\omega))\, dP(\omega)}_{E[f(X)]} \equiv \int_{\mathbb{R}} f(t)\, d\delta_c(t) = f(c)$$

by Change of Variable Thm

Note: The mapping $f \longmapsto E[f(X)]$ is known as an evaluation map. Why?

Because

$E[f(X)] = \underset{\text{eval. } f \text{ at } c}{f(c)}$

considered earlier on

Ex 2: Suppose RV X has the Poisson $(\lambda = 5)$ distribution.

$$P(X \in A) = \sum_{k \in A} \frac{e^{-5}}{5^k k!} \Rightarrow \boxed{\mathcal{L}(X) = \sum_{k=0}^{\infty} \left(\frac{e^{-5}}{5^k k!} \right) \delta_k}$$

distribution (law) of X

This distribution is a convex combination of point masses.

Then $$E[f(X)] = \sum_{k=0}^{\infty} f(k) \left(\frac{e^{-5}}{5^k k!} \right) \delta_k$$ for any function $f : \mathbb{R} \to \mathbb{R}$.

Lec 14 (1a)

SKIP

Prop 6.2.1: Suppose $\mu = \sum_i \beta_i \mu_i$ where $\{\mu_i\}$ are probability distributions and $\{\beta_i\}$ are non-negative constants (summing to 1, if we want μ to also be a probability dist'n).
Then for Borel-measurable functions $f : \mathbb{R} \to \mathbb{R}$,

$$\int f \, d\mu = \sum_i \beta_i \int f \, d\mu_i$$

provided either side is well-defined.

Ex 3: Suppose RV X has the Normal $(0,1)$ dist'n.
$(X \sim N(0,1))$

Distribution of X: $$\boxed{\mu_N(B) = \int_{-\infty}^{\infty} \phi(t) \, \mathbb{1}_B(t) \, d\lambda(t)},$$ $B \in \mathcal{B}$ (Borel set)

λ = Lebesgue meas on $\mathbb{R}$

$f(x)$

$\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$