Distributions of Random Variables

(ref: Rosenthal §6, Billingsley §20)

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

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

Notation: If $\mu$ is the distribution of $X$, then

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

def: The cumulative distribution function of a RV $X$

by $F_X(x) = P(X \leq x)$ for $x \in \mathbb{R}$. 
Properties of CDF

- By continuity of probabilities, \( F_x \) is right-continuous:
  \[
  \text{i.e., if } \{ x_n \} \nearrow x \text{ then } F_x(x_n) \to F_x(x).
  \]

- \( F_x \) 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 \( X \perp Y \) be two RVs (possibly defined on different prob. spaces). Then
\[
L_X(x) = L_Y(y) \iff F_x(x) = F_y(y) \quad \text{for } x \in \mathbb{R}.
\]

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

\[\text{Thm [Change of Variable Theorem]}:\]
Given a probability space \((\Omega, \mathcal{F}, P)\), let \( X \) be a RV having distribution \( \mu \). Then for any Borel-measurable function \( f: \mathbb{R} \to \mathbb{R} \), we have
\[
\int_{\Omega} f(X(w)) \, dP(w) = \int_{-\infty}^{\infty} f(t) \, d\mu(t)
\]
\[\text{Note: } \int_{\Omega} \text{ or } P(dw) \quad \text{or } \mu(dt) \]
i.e. \( E_P[f(X)] = E_{\mu}[f] \) provided that either side is defined.
In words,
the exp. value of RV \( f(X) \) w.r.t. the prob. measure \( P \)
on \( \Omega \) is equal to the exp. value of the function \( f \)
w.r.t. the measure \( \mu \) on \( \mathbb{R} \).

**Cor:** Let \( X \) & \( Y \) be 2 RVs (possibly defined on different prob. spaces). Then \( L(X) = L(Y) \) 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 \( X \neq Y \) are RVs s.t. \( P(X = Y) = 1 \), then

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

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

**Proof [Change of Var. Thm]:** First suppose that \( f = \mathbb{1}_B \) for \( B \in \mathcal{B} \).

Then

\[
\int_{\Omega} f(X(w)) \, dP(w) = \int_{\Omega} \mathbb{1}_{\{X(w) \in B\}} \, dP(w) = P(X \in B),
\]

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. $\{f_n\} \uparrow f$. We know (by above argument) that (6.1.2) holds when $f$ is replaced by $f_n$. Let $n \to \infty$ and 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 fun
- non-neg. general fun
- general fun


Examples of Distributions

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

The distribution of $X$ is the point mass $\delta_c$

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 $\mathbb{L}(X) = \delta_c$

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

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

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

by change of variable Thm

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

Because $E[f(X)] = f(c)$

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

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

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 $f: \mathbb{R} \rightarrow \mathbb{R}$.

Prop 6.2.1: Suppose $\mu = \sum_{i} \beta_i \mu_i$ where $\sum \mu_i$ are probability distributions and $\beta_i$ are non-negative constants (summing to 1, if we want $\mu$ to also be a probability dist'n).

Then for Borel-measurable functions $f: \mathbb{R} \rightarrow \mathbb{R}$,

$$\int f \, d\mu = \sum_{i} \beta_i \int f \, d\mu_i$$

provided either side is well-defined.

Ex 3: Suppose RV $X$ has the Normal (0,1) distribution.

$$(X \sim N(0,1))$$

Distribution of $X$:

$$\mu_X(B) = \int_{-\infty}^{\infty} \phi(t) \mathbb{1}_B(t) \, d\lambda(t) \quad \text{for} \quad B \in B$$

$\lambda = \text{Lebesgue meas}$ on $\mathbb{R}$

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