a Mahmood

Distribution of Random Variables

Given a random variable X defined on a probability triple $X(\Omega, \mathcal{F}, P)$, its distribution ~~is the~~ is the function μ defined on $\mathcal{B}$, the Borel subset of $\mathbb{R}$, denoted as

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

If μ is the law of random variable, then $(\mathbb{R}, \mathcal{B}, \mu)$ is a valid probability triple. We shall sometimes write μ as $\mathcal{L}(X)$ or as $P(X^{-1})$. We can also write $X \sim \mu$ to indicate that μ is the distribution of X.

We define cummulative distribution function of a random variable X by $F_X(x) = P(X \leq x)$, for $x \in \mathbb{R}$. By continuity of probabilities, the function F_X is right-continuous, i.e. if $\{x_n\} \searrow x$ then $F_X(x_n) \to F_X(x)$. It is also clearly a non-decreasing function of x with $\lim\limits_{x \to \infty} F_X(x) = 1$ and $\lim\limits_{x \to -\infty} F_X(x) = 0$.

Theorem 1: (Change of Variable Theorem) Given a probability triple $(\Omega, \mathcal{F}, P)$ let X be a random variable having distribution μ. Then for any Borel-measurable function $f: \mathbb{R} \to \mathbb{R}$, we have

$$\int_{\Omega} f(X(\omega))\, P(d\omega) = \int_{-\infty}^{\infty} f(t)\, \mu(dt)$$

Proof: In words, the expected value of the random variable $f(X)$ with respect to the probability measure P on Ω is equal to the expected value of of the function f with respect to the measure μ on $\mathbb{R}$.

First, suppose that $f = 1_B$ is an indicator function of a Borel set $B \subseteq \mathbb{R}$. Then $\int_\Omega f(X(\omega))\, P(d\omega) = \int_\Omega 1_{[X(\omega) \in B]}\, P(d\omega)$ $= P(X \in B)$, while $\int_{-\infty}^{\infty} f(t)\mu(dt) = \int_{-\infty}^{\infty} 1_{[t \in B]}\, \mu(dt) = \mu(B) = P(X \in B)$

So the equality holds when f is an indicator function.

Now suppose that f is a non-negative simple function, the f is a finite positive linear combination of indicator functions. But since both sides in Th 1 are linear functions of f, the equality will hold in this case as well.

Next we suppose that f is a general non-negative Borel-measurable function. Then as discussed earlier, we can find a sequence $\{f_n\}$ of non-negative simple functions such that $\{f_n\} \to f$. Then using the Monotone Convergence Theorem, we say that Th 1 also holds for f as well.

~~Next~~ Finally, for any general Borel-measurable f, we write $\boxed{f = f^+ - f^-}$ Since Th1 holds for f^+ and f^- separately, and since it is linear, therefore it is also true for f.

Theorem 2: Let X and Y be two random variables. Then $\mathcal{L}(X) = \mathcal{L}(Y)$ if and only if $E[f(X)] = E[f(Y)]$ for all Borel-measurable $f: \mathbb{R} \to \mathbb{R}$ for which either expectation is well-defined.

Proof: If $\mathcal{L}(X) = \mathcal{L}(Y) = \mu$ then straightforward from Th 1 we know that $E[f(X)] = E[f(Y)] = \int_{\mathbb{R}} f \, d\mu$.

Conversely, if $E[f(X)] = E[f(Y)]$ for all Borel-measurable $f: \mathbb{R} \to \mathbb{R}$ then setting $f = 1_B$ shows that $P(X \in B) = P(Y \in B)$ for all Borel $B \subseteq \mathbb{R} \Rightarrow \mathcal{L}(X) = \mathcal{L}(Y)$.

Theorem 3: If X and Y are random variables with $P(X = Y) = 1$ then $E[f(X)] = E[f(Y)]$ for all Borel-measurable $f: \mathbb{R} \to \mathbb{R}$ for which either expectation is well defined.

Proof: From $P(X = Y) = 1$ we can directly say that $\mathcal{L}(X) = \mathcal{L}(Y)$. Then letting $\mu = \mathcal{L}(X) = \mathcal{L}(Y)$ we have Theorem 1, $E[f(X)] = E[f(Y)] = \int_{\mathbb{R}} f \, d\mu$.

Theorem 4: Suppose that $\mu = \sum_i B_i \mu_i$, where $\{\mu_i\}$ are probability distributions and $\{B_i\}$ are non-negative constants (they can be summing to 1). Then for Borel-measurable functions $f: \mathbb{R} \to \mathbb{R}$,

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

provided that either side is well defined.

Proof: As in the proof of Th 1, it suffices by linearity and M.C.T. to check the equation $f = 1_B$ when it is an indicator function of a Borel set B. But in this case the result is evident, ~~#~~ since $\mu(B) = \sum_i \beta_i \mu_i(B)$. $\square$

Def: Given any Borel-measurable function called a density function f such that $f \geq 0$ and $\boxed{\int_{-\infty}^{\infty} f(t)\lambda(dt) = 1}$, where λ is a Lebesgue measure on $\mathbb{R}$. Then we define the dist. μ by,

$$\mu(B) = \int_{-\infty}^{\infty} f(t) 1_B(t) \lambda(dt),$$

We shall also write this as $\mu(B) = \int_B f(t)\lambda(dt)$, or also as, $\mu(dt) = f(t)\lambda(dt) \Longrightarrow \boxed{\int_B \mu(dt) = \int_B f(t)\lambda(dt)}$ We shall also write this as $\frac{d\mu}{d\lambda} = f$, we shall say that μ is absolutely continuous with respect to λ.

Theorem 5: Suppose that μ has a density f with respect to λ. Then for any Borel-measurable function $g : \mathbb{R} \to \mathbb{R}$

$$E_\mu(g) = \int_{-\infty}^{\infty} g(t)\mu(dt) = \int_{-\infty}^{\infty} g(t) f(t) \lambda(dt)$$

Proof: In words, to compute the integral of a function with respect to μ, it suffices to compute the integral of the function times density with respect to λ.

We will take a similar strategy like the one used to prove Th 1. We will first check for the case when $\boxed{g = 1_B}$ is an indicator function of the Borel set B. But in that case,

$$\int g(t)\mu(dt) = \int 1_B(t)\mu(dt) = \mu(B), \text{ while } \int g(t)f(t)\lambda(dt)$$

$$= \int 1_B(t) f(t)\lambda(dt) = \mu(B) \text{ by definition of } f \text{ and } \lambda. \square$$

For further illustrations, let $X \sim N(0,1)$. Then we have

$$E(X) = 0, \quad E(X^2) = 1, \quad E(X^4) = 3 \ldots$$

Suppose that $\mathcal{L}(X) = \frac{1}{4}\delta_1 + \frac{1}{4}\delta_2 + \frac{1}{2}\mu_N$ (δ_c is point mass c).

Then $E(X) = \frac{1}{4}(1) + \frac{1}{4}(2) + \frac{1}{2}(0) = \frac{3}{4}$

$$E(X^2) = \frac{1}{4}(1) + \frac{1}{4}(4) + \frac{1}{2}(1) = \frac{7}{4}$$

$$\mathrm{Var}(X) = E(X^2) - E(X)^2 = \frac{7}{4} - \left(\frac{3}{4}\right)^2 = \frac{19}{16}.$$