PROBABILITY THEORY

MATH 154

Unit 6: Distribution Functions

6.1. The **cumulative distribution function** of a random variable X is defined as

$$F_X(s) = \mu_X((-\infty, s]) = \mathrm{P}[X \leq s] \ ,$$

where μ_X is the law of X. It is often abbreviated as **CDF**.

6.2. If $F_X(s)$ is differentiable, it defines the **probability density function** $f_X(s) = F'_X(s)$, abbreviated **PDF**. In that case the law is $\mu_X = f_X(s)ds$ meaning $\mu([a,b)) = \int_a^b f_X(x) \ dx$ and $F_X(s) = \int_{-\infty}^s f_X(x) \ dx$. Random variables with a PDF are called **continuous random variables**.

6.3. In general, the function F_X is monotone increasing and satisfies $F_X(-\infty) = 0$ and $F_X(+\infty) = 1$. It does not have to be smooth. For example, if $X(x) = (-1)^x$ is the random variable on $(\Omega = \{1,2,3,4,5,6\}, \mathcal{A} = 2^\Omega, \mathrm{P}(\{x\}) = 1/6)$ then F_X jumps at the values -1 and 1 and is constant elsewhere. The law of X is the probability measure which is supported on $\{-1, 1\}$ and has weights $1/2$ on both points.

6.4. The distribution function F is useful: in order to get random variables with a distribution function F, just take a random variable Y with uniform distribution on $[0,1]$. Now, $X = F^{-1}(Y)$ has the distribution function F because $\{X \in [a,b] \ \} = \{F^{-1}(Y) \in [a,b] \ \} = \{Y \in F([a,b]) \ \} = \{Y \in [F(a), F(b)]\} = F(b) - F(a)$.

6.5. Example. Assume we want to generate a random variable with Cauchy distribution with PDF $f(x) = F'(x) = (\frac{1}{\pi}/(1+x^2)$. Integrating gives $F(x) = \frac{1}{2} + \arctan(x)/\pi$ and $F^{-1}(x) = \tan(\pi x - \pi/2) = \cot(\pi x)$. We can therefore compute Cauchy distributed random variables by evaluating $\cot(\pi x)$ with uniformly distributed random variables in $[0,1]$.

6.6. A measure μ on $(\mathbb{R}, \mathcal{B})$ is **absolutely continuous** (ac) with respect to the **Lebesgue measure** $\lambda = dx$ if $\lambda(A) = 0 \Rightarrow \mu(A) = 0, \forall A \in \mathcal{B}$. If $\mu = f d\lambda$ meaning that there exists a non-negative measurable function f satisfying $F(x) = \mu([-\infty, x]) = \int_{-\infty}^x f(x) \ dx$. A measure is **pure point** (pp) if there exists a finite or countable set of real numbers x_n and a sequence of positive numbers $p_n, \sum_n p_n = 1$ with $F(x) = \mu([-\infty, x]) = \sum_{n, x_n \leq x} p_n$. Finally, μ is called **singular continuous** (sc) if μ is continuous $(\mu(\{x\}) = 0$ for all x and $\mu(S) = 1$ for some set S of zero Lebesgue measure. [1]

[1] Many textbooks simply use **continuous** for (ac).

6.7. Nomenclature for μ goes over to CDF's F or random variables X. A (pp) measure μ could be dense on some intervals. It is supported on **atoms**, points x with $\mu(\{x\}) > 0$. For (ac) measures, there is a density function f meaning $\mu = f\ dx$ and $\int_{-\infty}^{\infty} f\ dx = 1$. The existence of this **Radon-Nykodym derivative** is to take the supremum over all non-negative functions with the property $\int_A f d\lambda \leq \mu(A)$ and $\int_{\mathbb{R}} f d\lambda = 1$. A (sc) measure μ has no atoms and satisfies $\mu(S) = 1$ for some $S \in \mathcal{B}$ with $\lambda(S) = 0$.

6.8. These three classes are fundamental:

Theorem 1 (Lebesgue decomposition theorem)**.** *Every probability measure μ on $(\mathbb{R}, \mathcal{B})$ can be decomposed uniquely into $\mu = \mu_{pp} + \mu_{ac} + \mu_{sc}$, where μ_{pp} is pure point, μ_{sc} is singular continuous and μ_{ac} is absolutely continuous.*

Proof. (i) First verify that any measure μ can be decomposed as $\mu = \mu_{ac} + \mu_s$, where μ_{ac} is absolutely continuous with respect to λ and μ_s is singular. The decomposition is unique: $\mu = \mu_{ac}^{(1)} + \mu_s^{(1)} = \mu_{ac}^{(2)} + \mu_s^{(2)}$ implies that $\mu_{ac}^{(1)} - \mu_{ac}^{(2)} = \mu_s^{(2)} - \mu_s^{(1)}$ is both absolutely continuous and singular with respect to μ which is only possible, if they are zero. For existence define $c = \sup_{A \in \mathcal{A}, \lambda(A)=0} \mu(A)$. If $c = 0$, then μ is absolutely continuous and we are done. If $c > 0$, take an increasing sequence $A_n \in \mathcal{B}$ with $\mu(A_n) \to c$. Define $A = \bigcup_{n \geq 1} A_n$ and μ_s as $\mu_s(B) = \mu(A \cap B)$. (ii) To split the singular part $\mu_s = \mu - \mu_{ac}$ into a singular continuous and pure point part, we again have uniqueness because $\mu_s = \mu_{sc}^{(1)} + \mu_{pp}^{(1)} = \mu_{sc}^{(2)} + \mu_{pp}^{(2)}$ implies that $\nu = \mu_{sc}^{(1)} - \mu_{sc}^{(2)} = \mu_{pp}^{(2)} - \mu_{pp}^{(1)}$ are both singular continuous and pure point which implies that $\nu = 0$. To get existence, define the finite or countable set $A = \{\omega \mid \mu(\omega) > 0\ \}$ and define $\mu_{pp}(B) = \mu(A \cap B)$ and $\mu_{sc} = \mu - \mu_{pp}$. ∎

6.9. Examples of absolutely continuous distributions

- The **Normal distribution** $N(m, \sigma^2)$ on $\mathbb{R}$ has PDF $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$.
- The **Cauchy distribution** on $\mathbb{R}$ has $f(x) = \frac{1}{\pi} \frac{b}{b^2 + (x-m)^2}$.
- The **exponential distribution** on $\mathbb{R}^+$ has $f(x) = \lambda e^{-\lambda x}$.

6.10. Example of a singular continuous distribution

- The Cantor distribution on $[0, 1]$ supported on the Cantor middle third. Its CDF is called the **Cantor staircase**.

6.11. Examples of pure point distributions:

- The **Binomial distribution** on $\{1, \dots, n\ \}$ $\mathrm{P}[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$.
- The **Poisson distribution** on $\mathbb{N}$ has $\mathrm{P}[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}$
- The **geometric distribution** on $\mathbb{N} = \{0, 1, 2, 3, \dots\ \}$ has $\mathrm{P}[X = k] = p(1-p)^k$.

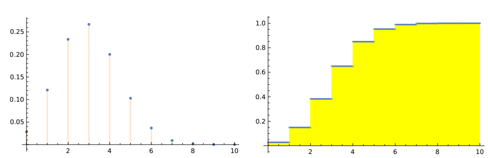

FIGURE 1. The law and CDF of a Binomial distribution.