Math 617: Day 46

Probability

In probability theory, we talk about a sample space $(\Omega, \mathcal{B}, P)$, which is just a finite measure space where the measure is normalized s.t. $P(\Omega) = 1$.

In this language an event is just a measurable set & the probability of the event is just the measure of the set.

Typically, if E is measurable, we write things like $P(x \in E)$ rather than $P(E)$ to emphasize the sense in which this is an event.

Def: Let $(R, \mathcal{R})$ be a measurable space. A random variable on $(\Omega, \mathcal{B}, P)$ is a measurable map X from the sample space to R. In other words, $X : \Omega \to R$ with $X^{-1}(S)$ an event for every $S \in \mathcal{R}$.

Ex: · $(R, \mathcal{R})$, where R is finite or countable & $\mathcal{R} = \mathcal{P}(R)$. Then $X : \Omega \to R$ is a discrete random variable. If $R = \{0,1\}$, then X is Boolean.

· $(R, \mathcal{R}) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, in which case X is a real-valued random variable

· If X_i is R_i-valued for $i = 1, 2$, then we have the joint random variable (X_1, X_2) taking values in $R_1 \times R_2$ given by $(X_1, X_2)(\omega) = (X_1(\omega), X_2(\omega))$.

Given a random variable X taking values in R, we define the distribution of X, denoted P_X, to be the probability measure on $(R, \mathcal{R})$ given by $P_X(S) := P(X \in S)$.

In other words, $P_X = X_* P$ is the pushforward of P by X.

Ex: For discrete random variables, $P_X(S) = \sum_{s \in S} p_s$, where $p_s = P(X = s)$ are nonnegative reals that sum to 1. For example,

· Discrete uniform distribution, when $|R| < \infty$ & $p_x = \frac{1}{|R|}$ $\forall x \in R$.

· Bernoulli distribution when $R = \{0,1\}$, $p_1 = p$, $p_0 = 1 - p$ for some $p \in [0,1]$.

· Geometric distribution when $R = \{0, 1, 2, \ldots\}$ & $p_k = (1-p)^k p$ for some $p \in [0,1]$.

· Poisson distribution when $R = \{0, 1, 2, \ldots\}$ & $p_k = \frac{\lambda^k e^{-\lambda}}{k!}$ for some $\lambda > 0$.

We say a random variable is continuous if $P(X = x) = 0$ $\forall x \in R$. Given a reference measure m on R, X is absolutely continuous if $P(X \in S) = 0$ $\forall$ null sets S in R; i.e., $P_X \ll m$ & so $P_X = m_f$ & f is called the probability density function (pdf) of X.

When X is real-valued, we can also describe P_X in terms of the cumulative distribution function (cdf)

$$F_X(t) := P_X((-\infty, t]) = P(X \le t)$$

Of course, P_X is just the Lebesgue-Stieltjes measure of F_X. Moreover, if X is absolutely continuous (w.r.t. Lebesgue measure), then

$$\frac{d}{dt} F_X = f$$ a.e., where f is the pdf of X

Ex: · The uniform distribution on an interval I has pdf $f = \frac{1}{m(I)} \chi_I$.

· The Gaussian distribution w/ mean μ & variance σ^2 has pdf $f(x) := \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

Def: If X is a non-negative random variable, the expectation or expected value or mean of X is

$$E(X) := \int_0^\infty t \, dP_X(t).$$

Of course, by HW3 #6,

$$E(X) = \int_0^\infty t \, dP_X(t) = \int_0^\infty P_X(t \ge \lambda)\, d\lambda = \int_0^\infty P(X \in \{t : t \ge \lambda\})\, d\lambda = \int_0^\infty P(X \ge \lambda)\, d\lambda.$$

In turn, if X is real- or complex-valued & $E(|X|) < \infty$, then we can define

$$E(X) := \int_{\mathbb{R} \text{ or } \mathbb{C}} t \, dP_X(t).$$

Of course, since integrals are linear, we get linearity of expectation: $E(c_1 X_1 + \dots + c_n X_n) = c_1 E(X_1) + \dots + c_n E(X_n)$

If X is (say) real-valued & absolutely integrable & $F : \mathbb{R} \to [0, +\infty]$ is measurable, then we have the change of variables

$$E(F(X)) = \int_0^\infty t \, dP_{F \circ X}(t) = \int_0^\infty t \, d(F_* P_X)(t) = \int_0^\infty F(t)\, dP_X(t)$$

Similarly for $F : \mathbb{R} \to \mathbb{R}$ or $\mathbb{C}$ if $F(X)$ is absolutely integrable. The key examples of $E(F(X))$ are:

① the k^{th} moment $E(|X|^k) = \int |x|^k \, dP_x$

② The exponential moments $E(e^{tX})$ for real t

③ the characteristic function $E(e^{itX})$