Probability theory

Gabriel Ng

May 21, 2022

1 Measures and Inequalities

A measurable space is a pair $(E, \mathcal{E})$ where E is a set and $\mathcal{E}$ is a sigma algebra on E. A probability space is a measurable space and a positive measure $\mathbb{P}$ on our space such that $\mathbb{P}(E) = 1$. We denote such a space by $(\Omega, \mathcal{F}, \mathbb{P})$. The elements of $\mathcal{F}$ are called events.

If $(\Omega \mathcal{F}, \mathbb{P})$ and $(E, \mathcal{E})$ are probability and measurable spaces respectively, a function $f : \Omega \to E$ is called a random variable if it is measurable. It is a real random variable when $E = \mathbb{R}$ and $\mathcal{E} = B(R)$ the Borel sigma algebra. When $\int X \mathrm{d}\mathbb{P} < \infty$, we say X is integrable and denote this integral EX. Two random variables are equivalent if $\mathbb{P}(X = Y) = 1$.

The spaces L^p is defined by the set of f so that $(E|f|^p)^{1/p}$ is finite up to equivalence (equal a.s.). L^∞ is defined by the set of f such that $|f(x)|$ is finite aside from a subset of measure 0.

> **Proposition – Jensen's inequality**
> Let X be an integrable random variable on $\mathbb{R}^m$ and let $\Phi : \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ be a lower semi-continuous convex function. Then $\Phi(X)$ is lower semi-integrable and
>
> $$E\Phi(X) \geq \phi(EX).$$
>
> Furthermore, if Φ is strictly convex and $\Phi(EX) < \infty$ then the inequality is strict unless X takes only 1 value a.s.

Proof. X is integrable so that EX is finite. Since Φ is convex, there is a line through EX and $\Phi(EX)$ such that the graph of Φ lies above this line (this is determined by the lower and upper derivatives). Hence $\Phi(X) \geq \Phi(EX) + c(X - EX)$ so that $E\Phi(X) \geq \Phi(EX)$. Now, unless $\Phi(X) = \Phi(EX)$, then our second statement holds.

> **Proposition – Holder's inequality**
> Let Z, W be positive random variables and α, β positive real numbers such that $\alpha + \beta = 1$. Then
>
> $$E(Z^\alpha W^\beta) \leq (EZ)^\alpha (EW)^\beta.$$

Proof. Consider first the convex function $\exp(X)$. We know this is convex since it is a smooth function and its second derivative is greater than 0. Setting $a = \alpha \log\left(\frac{|Z|}{\|Z\|_\alpha}\right)$ and $b = \beta \log\left(\frac{|W|}{\|(\|W)_\beta}\right)$, we get using the definition of convexity that $\frac{|f||g|}{\|f\|_\alpha \|g\|_\beta} \leq \frac{|f|^\alpha}{\|f\|_\alpha^\alpha \alpha} + \frac{|g|^\beta}{\|g\|_\beta^\beta \beta}$ upon which we obtain our inequality after integrating.

> **Proposition – Minkowski's inequality**

Let X and Y be real random variables and $p \geq 1$. Then

$$\|X+Y\|_p \leq \|X\|_p + \|Y\|_p .$$

Proof. Use the Holder inequality and show that $(a+b)^p \leq 2^{p-1}a^p + 2^{p-1}b^p$.

It follows immediately from Minkowski's lemma that L^p is a vector space. Furthermore, Jensen's inequality on the convex function $|x|^{p/q}$ gives the inclusion of L^p into L^q whenever $p \leq q$.

2 Variance, Covariance and the distribution of a random variable

If X is a square integrable random variable $EX^2 < \infty$ or $X \in L^2$, its variance is the quantity

$$\operatorname{Var} X = E[(X-EX)^2] = EX^2 - (EX)^2.$$

We call the value EX^α the moment of order α (when it exists and is finite).

Proposition – Markov's inequality
For every $\delta > 0$ and $\beta > 0$, we have

$$\mathbb{P}(|X| \geq \delta) \leq \frac{E|X|^\beta}{\delta^\beta}.$$

Proof. Consider $A = \{\omega : |X(\omega)| \geq \delta\}$. Then $\chi_A \leq \chi_A \cdot (|X|^\beta/\delta^\beta)$ so that integrating gives our inequality.

Corollary – Chebyshev's inequality
Let $X \in L^2$. Then for every $\alpha > 0$

$$\mathbb{P}(|X-EX| \geq \alpha) \leq \frac{\operatorname{Var} X}{\alpha^2}.$$

We define the distribution of X as the measure on the pullback. More formally, given random variable $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{E})$, for $A \in \mathcal{E}$, the measure μ_X is defined by $\mu_X(A) = \mathbb{P}(X^{-1}(A))$. We can compute integrals under the image law via the following proposition:

Proposition
Let $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{E})$ be a random variable, μ_X its associated measure. Then a measurable function $f : (E, \mathcal{E}) \to (\mathbb{R}, B(R))$ is μ_X-integrable iff $f(X)$ is $\mathbb{P}$-integrable and we have

$$\int_E f \mathrm{d}\mu_X = \int_\Omega f \circ X \mathrm{d}\mathbb{P}.$$

Proof. We show it is first true for characteristic functions, then it is true for simple random variables by linearity and hence positive random variables and so for all random variables. Let $f = \chi_A$ for some $A \in \mathcal{E}$. Then by definition we have our statement. Our proposition follows.

Given two measures μ and ν, we say μ is absolutely continuous with respect to ν if $\mu(A) = 0$ implies $\nu(A) = 0$. We denote this by $\mu \gg \nu$.

Proposition – Radon-Nikodym theorem

> If $mu \gg \nu$ then there exists a random variable $f \geq 0$ such that
>
> $$\nu(A) = \int_A f \mathrm{d}\mu.$$

The proof of this theorem is relegated to the appendix.

3 Appendix

3.1 Radon-Nikodym