Want to check that expected value retains its basic properties: <u>order-preserving</u> & <u>linearity</u>.

[See Rosenthal for details]

ec 10 1a

Connection with the Integral

<u>def</u>: Let X be a random variable on a probability space $(\Omega, \mathcal{F}, P)$. The <u>expected value</u> of X is

$$\boxed{E[X] = \int_\Omega X(\omega)\, dP(\omega)}. \quad \text{or } P(d\omega) \text{ or } dP$$

(sometimes "Ω" is omitted above).

This is the Lebesgue integral of the measurable function X with respect to the probability measure P.

<u>Thm</u>: Let $(\Omega, \mathcal{F}, P)$ be Lebesgue measure on $[0,1]$. Let $X: [0,1] \to \mathbb{R}$ be a bounded function which is Riemann integrable. Then X is a random variable w.r.t. $(\Omega, \mathcal{F}, P)$ and $E[X] = \int_0^1 X(t)\, dt$.

(Special case!)

There are many functions X which are NOT Riemann integrable, but are RVs w.r.t. Lebesgue measure, & thus have well-defined exp. values.

Example : Let $X(\omega) = \begin{cases} 0 & \text{if } \omega \in \mathbb{Q} \text{ (rational)} \\ 1 & \text{if } \omega \notin \mathbb{Q} \text{ (irrational)} \end{cases}$

i.e. $X = \mathbb{1}_{\mathbb{Q}^c}$ where $\mathbb{Q}^c$ = irrational numbers

Then X is not Riemann integrable, but we still have

$$E[X] = \int_{\mathbb{R}} X \, dP = 1 \quad \text{well-defined}$$

Understanding the Integral-Expectation Connection

X is a RV on a probability space $(\Omega, \mathcal{F}, P)$

$$E[X] = \int_{\Omega} X(\omega) \, dP(\omega) = \int_{\Omega} X \, dP$$

shorthand notation

Lebesgue integral

If $E[|X|] < \infty$, then X is Lebesgue-integrable.

Theorem: Let X be a RV on prob. space $(\Omega, \mathcal{F}, P)$.

Then for each Borel-measurable function g on $\mathbb{R}$,

we have

$$\boxed{E[g(X)] = \int_{\mathbb{R}} g(x)\, d\mu_X(x)}$$

M & W notation (from 713)

where μ_X is the "probability distribution" of X.

Recall: μ_X is defined by

$$\mu_X(B) = P(X \in B) \text{ for } B \in \mathcal{B}$$

Borel σ-algebra

e.g. X is an absolutely continuous RV if

$\exists$ non-neg. Borel-meas. function f_X (PDF of X)

s.t. $\mu_X(B) = \int_B f_X \, d\lambda \quad \forall B \in \mathcal{B}$.

Alternate def:

$$\boxed{E[g(X)] = \int_{\mathbb{R}} g(x)\, dF(x)}$$

where F is CDF of X.

Given a distribution function F, there is a unique probability function P_F s.t. $P_F((a,b]) = F(b) - F(a)$

$= \mu_X((a,b])$

from notation above

Inequalities and Convergence

(ref: §5 Rosenthal)

Markov's Inequality: If X is a *non-negative* random variable, then for all $\alpha > 0$,

$$\boxed{P(X \geq \alpha) \leq \frac{E[X]}{\alpha}}.$$

Pf: Define a new RV Z by

$$Z(\omega) = \begin{cases} \alpha, & X(\omega) \geq \alpha \\ 0, & X(\omega) < \alpha \end{cases}$$

Then $Z \leq X$ so it follows that $E[Z] \leq E[X]$ (order-preserving property). Compute

$$\begin{aligned} E[Z] &= \alpha \cdot P(X \geq \alpha) + 0 \cdot P(X < \alpha) \\ &= \alpha P(X \geq \alpha) \end{aligned}$$

$$\Rightarrow P(X \geq \alpha) = E[X]/\alpha.$$

Chebychev's Inequality: Let Y be an arbitrary RV with finite mean μ_Y. Then $\forall\, \alpha > 0$,

$$\boxed{P(|Y-\mu_Y| \geq \alpha) \leq \frac{Var(Y)}{\alpha^2}}$$

applies more generally.

Pf: Set $X = (Y-\mu_Y)^2$. Then X is a non-neg. RV.

By Markov's Inequality,

$$P(|Y-\mu_Y| \geq \alpha) = P(X \geq \alpha^2) \leq \frac{E[X]}{\alpha^2} = \frac{Var(Y)}{\alpha^2}.$$

We will use these 2 inequalities extensively, including to prove the laws of large numbers (see below).

Two other sometimes useful inequalities:

Cauchy-Schwarz Inequality:

Let X & Y be RVs with $E[X^2] < \infty$ and $E[Y^2] < \infty$.

Then $$\boxed{E[|XY|] \leq \sqrt{E[X^2]\,E[Y^2]}}.$$

Pf: Let $Z = \frac{|X|}{\sqrt{E[X^2]}}$ and $W = \frac{|Y|}{\sqrt{E[Y^2]}}$, so that

$$E[Z^2] = E[W^2] = 1. \quad \left(\text{since } E[Z^2] = E\left[\frac{|X|^2}{E[X^2]}\right] = \frac{E[X^2]}{E[X^2]} = 1 \right)$$

Then, $0 \le E[(Z-W)^2] = E[Z^2 + W^2 - 2ZW]$

$$= E[Z^2] + E[W^2] - 2E[ZW]$$

$$= 1 + 1 - 2E[ZW]$$

$$\Rightarrow \quad 2E[ZW] \le 2$$

$$\Rightarrow \quad E[ZW] \le 1$$

Thus, $E[ZW] = E\left[\frac{|X|}{\sqrt{E[X^2]}} \cdot \frac{|Y|}{\sqrt{E[Y^2]}}\right] = \frac{E[|XY|]}{\sqrt{E[X^2]E[Y^2]}} \le 1$

$$\Rightarrow \quad E[|XY|] \le \sqrt{E[X^2]\,E[Y^2]}. \quad \blacksquare$$

Jensen's Inequality:

Let X be a RV with finite mean and let $\phi: \mathbb{R} \to \mathbb{R}$ be a convex function, i.e. a function s.t.

$$\lambda\phi(x_1) + (1-\lambda)\phi(x_2) \ge \phi(\lambda x_1 + (1-\lambda)x_2) \text{ for } x_1, x_2, \lambda \in \mathbb{R}$$

and $0 \le \lambda \le 1$.

Then $\boxed{E[\phi(X)] \ge \phi(E[X])}$.

Pf: Since ϕ is convex, we can find a linear function $g(x) = ax + b$ which lies entirely below the graph of ϕ but which touches it at the point $x = E[X]$, i.e.

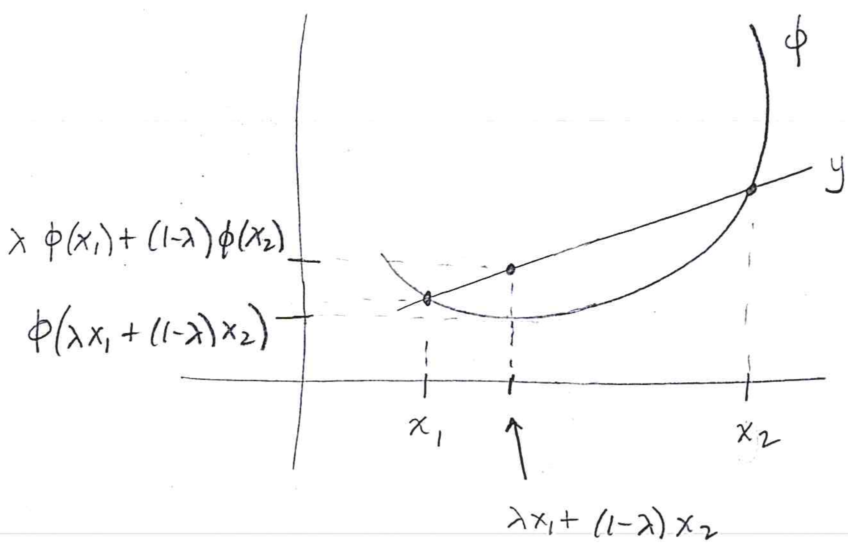
s.t. $g(x) \le \phi(x) \ \forall x \in \mathbb{R}$ and $g(E[X]) = \phi(E[X])$.

Then $E[\phi(X)] \geq E[g(X)]$ order-preserving of $E(\cdot)$

$$= E[aX+b]$$

$$= a E[X] + b$$

$$= g(E[X]) = \phi(E[X]). \quad \blacksquare$$



convex function:
if line segment b/t any 2 points on ϕ lies above or on the graph of ϕ.

Convergence of Random Variables (ref: R §5.2)

If Z and $Z_1, Z_2, \ldots$ are random variables defined on a probability space $(\Omega, \mathcal{F}, P)$, what does it mean to say that $\{Z_n\}$ converges to Z as $n \to \infty$?

Different notions of convergence:

- pointwise
- almost surely (with probability 1)
- in probability

Q. Which ones are strong, weak?

def: Pointwise convergence: $\lim_{n\to\infty} Z_n(\omega) = Z(\omega) \ \forall \omega.$

def: Convergence in distribution: $\lim_{n\to\infty} F_{X_n}(\omega) = F(\omega) \ \forall \omega$

Point wise convergence of CDFs

OR w.p. 1

def: Convergence almost surely (a.s.) : $Z_n \xrightarrow{a.s.} Z$

$$P\left(\lim_{n\to\infty} Z_n = Z\right) = 1 \qquad \left(\text{OR } P\left(\lim_{n\to\infty} |X_n - X| \le \varepsilon\right) = 1 \quad \forall \varepsilon > 0\right)$$

i.e. $P\left(\{\omega \in \Omega : \lim_{n\to\infty} Z_n(\omega) = Z(\omega)\}\right) = 1$

Lemma: Let Z and $Z_1, Z_2, \ldots$ be RVs. Suppose for each $\varepsilon > 0$, we have $P(|Z_n - Z| \geq \varepsilon \text{ i.o.}) = 0$.

Then $P(Z_n \to Z) = 1$, i.e. $Z_n \xrightarrow{a.s.} Z$.

Combine this Lemma with Borel-Cantelli Lemmas:

Corollary: Let Z and $Z_1, Z_2, \ldots$ be RVs. Suppose for each $\varepsilon > 0$, we have $\sum_n P(|Z_n - Z| \geq \varepsilon) < \infty$.

Then $P(Z_n \to Z) = 1$, i.e. $Z_n \xrightarrow{a.s.} Z$.

def: Convergence in Probability $\quad X_n \xrightarrow{P} X$

If $\forall \varepsilon > 0$, $\lim_{n\to\infty} P(|Z_n - Z| \geq \varepsilon) = 0$.

Note: Convergence a.s. $\Rightarrow$ convergence in probability

but converse NOT true.