Recall definition of expected value for simple RV:

X is simple RV (finite range), $X = \sum_{i=1}^{n} x_i \mathbb{1}_{A_i}$ where $\{A_i\}$ is finite partition of Ω

$$E[X] = \sum_{i=1}^{n} x_i P(A_i) \qquad A_i = \{\omega \in \Omega : X(\omega) = x_i\}$$

[Alternate form: $E[X] = \sum_{x} x \, P(X = x)$]

Remark: If $X = \sum_{i=1}^{n} x_i \mathbb{1}_{A_i}$ with $\{A_i\}$ a finite partition of Ω, and if $f : \mathbb{R} \to \mathbb{R}$ is any function, then
$f(X) = \sum_{i=1}^{n} f(x_i) \mathbb{1}_{A_i}$ is also a simple RV with

$$E[f(X)] = \sum_{i=1}^{n} f(x_i) P(A_i).$$

In particular, if $f(x) = (x - \mu_X)^2$, then we get the variance of X, defined by

$$\boxed{\mathrm{Var}(X) = E[(X - \mu_X)^2]} \quad \text{where } \mu_X = E[X].$$

Clearly $Var(X) \geq 0$. Expanding the square & using linearity,

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

Properties

- $Var(X) \leq E[X^2]$ (equal if mean of X is 0)

- $Var(aX+b) = a^2 \, Var(X)$

- $Var(X+Y) = Var(X) + Var(Y) + 2\,Cov(X,Y)$

  where $Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$

  $$= E[XY] - E[X]E[Y]$$

  Note: If X & Y are independent, then $Cov(X,Y) = 0$.

  (and hence $Var(X+Y) = Var(X) + Var(Y)$).

- More generally,

  $$Var\left(\sum_i X_i\right) = \sum_i Var(X_i) + 2\sum_{i<j} Cov(X_i, X_j)$$

So, $Var(X_1 + \cdots + X_n) = Var(X_1) + \cdots + Var(X_n)$ if $\{X_i\}$ independent.

- If $Var(X) > 0$ & $Var(Y) > 0$, then the <u>correlation</u> between X & Y is given by

$$\boxed{Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)\,Var(Y)}}}$$

← normalized covariance

Q. How to interpret this quantity?

max value = +1 → indicates perfect direct (increasing) linear relationship

min value = −1 → indicates perfect inverse (decreasing) linear relationship

value of 0 → indicates no relationship, uncorrelated

all other values in $(-1,1)$ denote degree of linear relshp.

<u>General non-negative random variables</u>

<u>def</u>: For a non-negative RV X, we define the <u>expected value</u> of X by

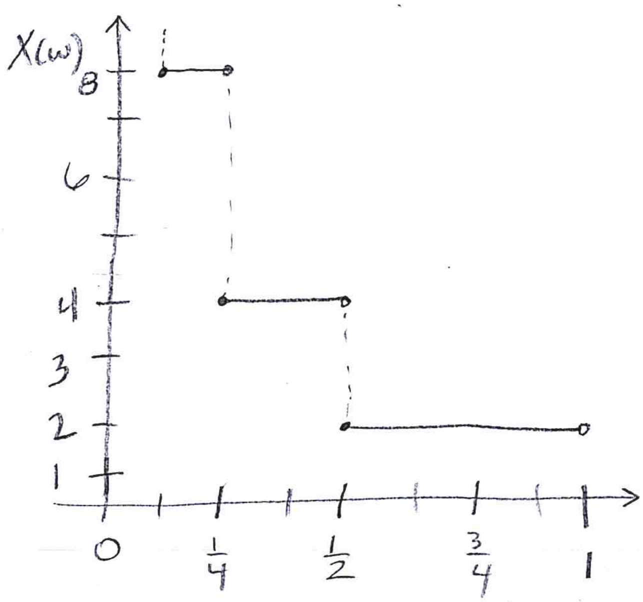
$$\boxed{E[X] = \sup\{E[Y] : Y \text{ simple}, Y \le X\}}$$

* This def'n (with a minor modification for negative values) will apply to all random variables
- discrete
- abs. continuous
- neither

<u>Note</u>: It is possible that $E[X]$ will be infinite.

<u>Example</u>: Suppose $(\Omega, \mathcal{F}, P)$ is Lebesgue measure on $[0,1]$. Define X by

$$X(\omega) = 2^n \text{ for } 2^{-n} \le \omega < 2^{-(n-1)}, \quad n \in \mathbb{N}$$



$X = 2^1$ for $2^{-1} \le \omega < 2^0$

$X = 2^2$ for $2^{-2} \le \omega < 2^{-1}$

$X = 2^3$ for $2^{-3} \le \omega < 2^{-2}$

$\vdots$

Then $E[X] \geq \sum_{i=1}^{N} 2^i 2^{-i} = \sum_{i=1}^{N} 1 = N$ for any $N \in \mathbb{N}$.

Hence, $\underline{E[X] = \infty}$.

$$\left(P(\omega \in A_i) \text{ where } A_i = \underbrace{[2^{-i}, 2^{-(i-1)})}_{\text{length of this interval is } 2^{-i}} \right)$$

Recall: For $k \in \mathbb{N}$, the k^{th} moment of a non-negative random variable X is $\boxed{E[X^k]}$, and is finite if $E[X^k] < \infty$ (in general, if $E[|X|^k] < \infty$).

Theorem: [Monotone Convergence Thm]

Suppose $X_1, X_2, \ldots$ are random variables with $E[X_1] > -\infty$ and $\{X_n\} \nearrow X$. Then X is a RV and

$$\lim_{n\to\infty} E[X_n] = E[X].$$

(Recall that $\underline{\{X_n\} \nearrow X}$ means: $X_1 \leq X_2 \leq \cdots$ and $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ for each $\omega \in \Omega$)

i.e. the sequence $\{X_n\}$ converges monotonically to X

* Use this to help prove linearity of $E(\cdot)$ for general non-neg RVs.

Remark: Since expected values are unchanged if we modify the RV values on sets of probability 0, we still have

$$\lim_{n\to\infty} E[X_n] = E[X] \text{ provided } \{X_n\} \nearrow X \ \underbrace{\text{almost surely}}_{\text{(a.s.)}}$$

i.e. on a subset of Ω having prob. 1.

Pf of MCT : Note that $\{X \le x\} = \bigcap_{n=1}^{\infty} \{X_n \le x\} \in \mathcal{F} \; \forall x \in \mathbb{R}$

so X is a RV. Monotonicity of X_n's gives us

$$E[X_1] \le E[X_2] \le \cdots \le E[X].$$

Hence, $\lim_{n\to\infty} E[X_n]$ exists (although it may be infinite if $E[X]=\infty$)

and $\lim_{n\to\infty} E[X_n] \le E[X]$. Lastly, we must show that

$\lim_{n\to\infty} E[X_n] \ge E[X]$, and then equality follows. If $E[X_1]=\infty$,

this is trivial, so let us assume $E[X_1]$ is finite. [Then, by

replacing X_n by $X_n - X_1$ & X by $X - X_1$, it suffices to

assume that X_n & X are non-negative.] By def of $E[X]$

for non-neg RV X, it suffices to show that $\lim_{n\to\infty} E[X_n] \ge E[Y]$

for any simple RV $Y \le X$. Write $Y = \sum_{i=1}^{m} v_i \mathbb{1}_{A_i}$, then

it suffices to prove that $\lim_{n\to\infty} E[X_n] \ge \sum_{i=1}^{m} v_i P(A_i)$

where $\{A_i\}$ is any finite partition of Ω with $v_i \le X(\omega) \; \forall \omega \in A_i$.

Let $\varepsilon > 0$ & set $A_{in} = \{\omega \in A_i : X_n(\omega) \ge v_i - \varepsilon\}$. Then

$\{A_{in}\} \nearrow A_i$ as $n \to \infty$. Moreover,

$$E[X_n] \ge \sum_{i=1}^{m} (v_i - \varepsilon) P(A_{in}).$$

As $n\to\infty$, by continuity of probabilities, this converges to

$$\sum_{i=1}^{m} (v_i - \varepsilon) P(A_i) = \sum_{i=1}^{m} v_i P(A_i) - \underbrace{\sum_{i=1}^{m} \varepsilon P(A_i)}_{\varepsilon \underbrace{\sum_{i=1}^{m} P(A_i)}_{1} = \varepsilon}$$

Thus, $\lim\limits_{n\to\infty} E[X_n] \geq \sum_{i=1}^{m} v_i P(A_i) - \varepsilon$

but this is true for any $\varepsilon > 0$, so $\lim\limits_{n\to\infty} E[X_n] \geq \sum_{i=1}^{m} v_i P(A_i)$ as required.

(by Prop A.3.1)

---

* Monotonicity assumption is necessary:

e.g. $(\Omega, \mathcal{F}, P)$ Lebesgue measure on $[0,1]$ & let

$X_n = n \mathbb{1}_{(0,\frac{1}{n})}$.

Then $X_n \to 0$ as $n \to \infty$ (since for each $w \in [0,1]$, $X_n(w) = 0$ for all $n \geq \frac{1}{w}$)

but $E[X_n] = 1 \quad \forall n \in \mathbb{N}$

$X_1 = 1$

$X_2 = \begin{cases} 2 & \text{if } w \in (0,\frac{1}{2}) \\ 0 & \text{if } w \in [\frac{1}{2},1) \end{cases}$

$X_3 = \begin{cases} 3 & \text{if } w \in (0,\frac{1}{3}) \\ 0 & \text{if } w \in [\frac{1}{3},1) \end{cases}$

$\vdots$

Arbitrary Random Variables

(ref § 4.3 Rosenthal, § 21 Billingsley)

Finally we consider RVs which may be neither simple nor non-negative.

For such a RV X, we can write $X = X^+ - X^-$

where
$$X^+(\omega) = \max(X(\omega), 0)$$
$$X^-(\omega) = \max(-X(\omega), 0).$$

Then both X^+ and X^- are non-negative so previous section results apply.

<u>def</u>: For a general RV X,

$$\boxed{E[X] = E[X^+] - E[X^-]}$$

<u>Note</u>: $E[X]$ is undefined if <u>both</u> $E[X^+]$ and $E[X^-]$ are infinite. However,

$$E[X^+] = \infty \text{ and } E[X^-] < \infty \Rightarrow E[X] = \infty$$
$$E[X^+] < \infty \text{ and } E[X^-] = \infty \Rightarrow E[X] = \infty$$

Clearly, $E[X^+] < \infty$ and $E[X^-] < \infty \Rightarrow E[X] < \infty$.

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

[See Rosenthal for details]

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

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

(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!)

Of course there are many functions X which are NOT Riem. integ but are RVs w.r.t. Lebesgue meas. & have well-defined $E[X]$ ...