Intuition. Atom = "not splittable".

Definition: atom

A is an **atom** if $P(A) > 0$ and $B \subseteq A \Rightarrow P(B) = 0$ or $P(B) = P(A)$.

Claim.

$$X(\omega) = \sum_{i=1}^{n} x_i I_{A_i}(\omega), A_1, \dots, A_n \text{ partition } \Omega.$$

1) $E[X] = \sum_{i=1}^{n} x_i P(A_i)$ is equivalent to the more familiar $E[X] = \sum_x xP(X = x)$.

Proof

$$\begin{aligned} \sum_x P(X = x) &= \sum_x P(\{\omega : X(\omega) = x\}) \\ &= \sum_{i=1}^{n} x_i P(\{\omega : X(\omega) = x_i\}) \text{ by simple r.v.} \\ &= \sum_{i=1}^{n} x_i P(A_i) \end{aligned}$$

□

2)

Definition: independent r.v.

R.v.s X and Y are independent if $\sigma(X)$ and $\sigma(Y)$ are independent. That is, given any $A \in \sigma(X)$ and $B \in \sigma(Y)$, $P(A \cap B) = P(A) \cdot P(B)$.

3)

Theorem

If X and Y are independent, then

$$E[XY] = E[X] \cdot E[Y].$$

Proof

Use independence to separate the joint indicator variable and thus the double sum.

We have $X(\omega) = \sum_{i=1}^n x_i I_{A_i}(\omega), Y(\omega) = \sum_{j=1}^m y_j I_{B_j}(\omega)$. Then

$$XY(\omega) = X(\omega) \cdot Y(\omega) = \sum_{i=1}^n \sum_{j=1}^m x_i y_i I_{A_i \cap B_j}^{(\omega)}.$$

Note $A_i = \{\omega : X(\omega) = x_i\} \in \sigma(X)$. Likewise $B_j \in \sigma(Y)$. Since X, Y are independent, $\sigma(X), \sigma(Y)$ are independent by definition. So

$$P(A_i \cap B_j) = P(A_i)P(B_j).$$

Therefore,

$$\begin{aligned} E[XY] &= \sum_{i=1}^n \sum_{j=1}^m x_i y_j P(A_i \cap B_j) \\ &= \sum_{i=1}^n x_i P(A_i) \sum_{j=1}^m y_j P(B_j) \\ &= E[X]E[Y] \end{aligned}$$

$\square$

4)

Theorem

If $X = \sum_{n=1}^\infty X_n$ *a.s.*, (*i.e.*: $P(\{\omega : X(\omega) = \sum_{n=1}^\infty X_n(\omega)\}) = 1$) and the partial sums of $\sum_{n=1}^\infty X_n$ are uniformly bounded, then

$$E[X] = E\left[\sum_{n=1}^\infty X_n\right] = \sum_{n=1}^\infty E[X_n].$$

Note. We can exchange sum and expected value in this case.

Proof

By linearly of $E[X]$ for partial sums. Let $S_n = \sum_{i=1}^n X_i, S =$

$\sum_{n=1}^{\infty} X_n$. Then $S_n \xrightarrow{a.s.} S$. By assumption, S_n are uniformly bounded. Thus, by previous theorem,

$$\lim_{n\to\infty} E[S_n] = E[S]$$

$$\lim_{n\to\infty} E\left[\sum_{i=1}^{n} X_i\right] = \lim_{n\to\infty} \sum_{i=1}^{n} E[X_i] = E\left[\sum_{n=1}^{\infty} X_n\right] \text{ by finite sum}$$

$$\sum_{n=1}^{\infty} E[X_n] = E\left[\sum_{n=1}^{\infty} X_n\right]$$

□

5)

Theorem

Let $g : \mathbb{R} \to \mathbb{R}$, then

$$g(X(\omega)) = \sum_{i=1}^{n} g(x_i) I_{A_i}(\omega)$$

and

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

Note. g might not be injective, "non-unique representation". And distinct representations of a r.v. gives us the same expectation. It is because if $A_i \cap B_j \neq \emptyset$, then $x_i = y_j$.

Proof

We want to show

$$\sum_{i=1}^{n} x_i P(A_i) = \sum_{j=1}^{m} y_i P(B_j)$$

and we do not assume that x_i, y_j are distinct respectively.

Since by assumption

$$X = \sum_{i=1}^{n} x_i I_{A_i} = \sum_{j=1}^{m} y_j I_{B_j}$$

Given $\omega \in A_i \cap B_j$, $x_i = y_j$, so we must have $x_i = y_j$ whenever $A_i \cap B_j \neq \emptyset$. Notice that since A_i, B_j form partitions of Ω,

$$P(A_i) = \sum_{j=1}^{m} P(A_i \cap B_j)$$
$$P(B_j) = \sum_{i=1}^{n} P(A_i \cap B_j)$$

Thus we have

$$\begin{aligned}
\sum_{i=1}^{n} x_i P(A_i) &= \sum_{i=1}^{n} x_i \sum_{j=1}^{m} P(A_i \cap B_j) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} x_i I_{A_i \cap B_j} \\
&= \sum_{j=1}^{m} \sum_{i=1}^{n} y_i I_{A_i \cap B_j} \\
&= \sum_{j=1}^{m} y_j P(B_j)
\end{aligned}$$

$\square$

6)

Corollary

If X, Y are independent r.v.s, then $E[g(X)h(Y)] = E[g(X)] \cdot E[h(Y)]$

Note. By (3) and (5).

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

Definition: variance

$$\sigma^2 = \mathrm{Var}[X] \coloneqq E[(X-\mu)^2].$$

1 Inequalities

Theorem: generalized Markov inequality

$X \geq 0$, $g(x)$ real-valued and non-negative, $c > 0$, then

$$P(g(X) \leq c) \leq \frac{E[g(X)]}{c}.$$

Proof

$$\begin{aligned}
E[g(X)] &= \sum_x g(x) \cdot P(A_i) \\
&= \sum_x g(x) P(X = x) \\
&= \sum_{\{x : g(x) \geq c\}} g(x) P(X = x) + \sum_{x : g(x) < c} g(x) P(X = x) \\
&\geq \sum_{\{x : g(x) \geq c\}} g(x) P(X = x) \text{ nonnegative 2nd term} \\
&\geq \sum_{\{x : g(x) \geq c\}} c \cdot P(X = x) \\
&= c \cdot P(g(X) \geq c)
\end{aligned}$$

$\square$