Computing expectation by conditioning. (§7.5).

- Recall computing probabilities by conditioning. Suppose $S = \bigcup_i F_i$ is a decomposition into mutually exclusive events $F_i$.

Then law of total probability (p. 20) states:

$$P(E) = \sum_i P(E \mid F_i) \cdot P(F_i)$$

- There is a similar:

\[ E[X] = \sum_i E[X \mid F_i] \cdot P(F_i) \]

Here $E[X \mid F]$ refers to the conditional expectation of $X$ given event $F$, defined as:

\[ E[X \mid F] = \sum_x x \cdot P\{X = x \mid F\} \]

(recall that $E[X] = \sum_x x \cdot P\{X = x\}$)

\[ \text{Proof of law of total Exp:} \]

\[ E[X] = \sum_x x \cdot P\{X = x\} \quad \text{(by def)} \]

\[ = \sum_x \left( \sum_i P\{X = x \mid F_i\} \cdot P(F_i) \right) \quad \text{(by law of Total Prob.)} \]

\[ = \sum_i \left( \sum_x x \cdot P\{X = x \mid F_i\} \right) \cdot P(F_i) = \sum_i E[X \mid F_i] \cdot P(F_i). \quad \text{QED.} \]

Ex. The average # of traffic accidents in Berkeley on a rainy day is 9.

It rains with prob. 0.2 in Berkeley.

What is the ave # of accidents per day in Berkeley?

\[ E[X] = E[X \mid \text{Rain}] \cdot P(\text{Rain}) + E[X \mid \text{Dry}] \cdot P(\text{Dry}) \quad \text{(by law of Total Exp)} \]

\[ = 9 \cdot 0.2 + 3 \cdot 0.8 = 4.2 \]
A miner is trapped in a mine with three doors. One door leads to a tunnel which takes the miner to safety in 2 hrs of travel. The other two doors are connected by a loop, which one can pass in 3 hrs. The miner is equally likely to choose any of the two doors at any time. ("memoryless"). What is the expected time it takes to reach safety?

**Condition on the miner's initial decision:**

\[
E[X] = E[X|G] P(G) + E[X|B] P(B)
\]

\[
= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{3} + 3 + E[X]
\]

Solving yields

\[
E[X] = \frac{2}{3} + \frac{2}{3} (3 + E[X])
\]

\[
E[X] = 8 \text{ hrs}
\]

---

4.8.1. Geometric distribution

"Time until first success":

**Def:** Consider an (infinite) sequence of indep. trials, with \( p \) = prob. of success in each trial.

let \( X = \# \) of the trial that is the first success

Then \( X \) is called a geometric r.v.; notation:

\( X \sim \text{Geom}(p) \)

**pmf:**

\[
p(k) = P\{X=k\} = (1-p)^{k-1} \cdot p, \quad k = 1, 2, ...
\]

- \( p_k \) = prob. that the first \( k-1 \) trials are failures
- \( p_k \) = prob. that the \( k^{th} \) trial is a success
Remark: Since \[ \sum_{k=1}^{\infty} p(k) = 1, \]

\[ \sum_{k=1}^{\infty} (1-p)^{k-1} p = 1 \implies \sum_{k=1}^{\infty} (1-p)^{k-1} = \frac{1}{p}. \] (Geometric series).

\[ E(X) = \begin{cases} \frac{1}{p} & \text{(after 1 failure, we start over again).} \\ \text{memoryless property} \end{cases} \]

\[ E[X] = \frac{1}{p} \]

Ex (refer to Ex. p. 48): How many doors on average will the minor open, until safety?

\[ Y \sim \text{Geom}(\frac{1}{2}) \implies E[Y] = 3 \text{ doors.} \]

(quality control).

Each time a notebook is manufactured, a fault may occur in the assembly line with prob. \( p = 0.05 \), which will cause defects in the laptop being manufactured and all further laptops.

The quality control department checks all manufactured laptops; a defect is found in a laptop with prob. \( q = 0.9 \).
When a defect is found, the laptop is disposed of, and the line gets repaired.

What is % of defective laptops produced?

\[ C = \text{"Good"}, \quad B = \text{"Bad"} \]

\[ X = \# \text{ of laptops until fault occurs}; \quad Y = \# \text{ of bad laptops until defect found}. \]

\[ X \sim \text{Geom}(p), \quad Y \sim \text{Geom}(q). \quad E[X] = \frac{1}{p}, \quad E[Y] = \frac{1}{q}. \]

Between two consecutive line repairs:

\[ \# \text{ of good laptops} = X-1, \quad \# \text{ of bad laptops} = Y-1 \quad (1 \text{ is disposed of}). \]

\[ E[L] = \frac{1}{p} - 1, \quad E[L] = \frac{1}{q} - 1. \]

\[ \text{Ans:} \quad \frac{\frac{1}{p} - 1}{\frac{1}{p} + \frac{1}{q} - 1} \approx 0.058 \quad (\approx 5.8\%). \]

Ex (for home): \( X \sim \text{Geom}(p) \implies \text{Var}(X) = \frac{1-p}{p^2} \) (Textbook Ex. 7.51).