Review: Random Variables

Sample space and events:

- Sample space $S$ consists of the set of all possible outcomes of an experiment
- An event $E$ is a subset of $S$, $E \subset S$
- $P(E)$ is probability of event $E$.

A random variable $X(\omega)$ maps from outcomes $\omega$ in the sample space $S$ to a real number.

- Often $\omega$ is dropped and just write $X$, leaving the $\omega$ as understood.
- The set of outcomes for which $X = x$ is $E_x = \{\omega | X(\omega) = x, \omega \in S\}$
- $P(X = x)$ is probability that random variable $X$ takes value $x$, probability mass function (PMF).
- $P(X = x)$ is probability that event $E_x$ occurs: $P(X = x) = P(E_x)$.
- $F(a) = P(X \leq a)$ is the cumulative distribution function (CDF)
- Indicator random variable $I_E$, $I_E = 1$ when event $E$ occurs and 0 otherwise. $P(I_E = 1) = P(E)$
Example

Roll a die:

- Sample space $S = \{1, 2, 3, 4, 5, 6\}$
- $E = \{1, 2\}$ is event that a 1 or a 2 is observed.
- $P(E)$ is probability of event $E$.
- Random variable $X(\omega) = 1$ when $\omega = 1$ or $\omega = 2$ and $X(\omega) = \text{when } \omega = 3, 4, 5, 6$.
- $E_1 = \{\omega | X(\omega) = 1, \omega \in S\} = \{1, 2\} = E$
- $E_0 = \{3, 4, 5, 6\} = E_1^c = E^c$.
- $P(X = 1) = P(E_1) = P(E) = \frac{2}{6}$.
- $P(X = 0) = P(E_0) = P(E^c) = 1 - P(E) = \frac{4}{6}$
Review: Random Variables

All of the rules for probabilities of events carry over to random variables using the fact that $P(X = x) = P(E_x)$

For two discrete random variables $X$ and $Y$ on same sample space $S$:

- $E_x = \{\omega \in S : X(\omega) = x\}$ is set of outcomes for which $X = x$,
  $E_y = \{\omega \in S : Y(\omega) = y\}$ is set of outcomes for which $Y = y$.
  $P(X = x) = P(E_x)$, $P(Y = y) = P(E_y)$

- $P(X = x \text{ and } Y = y) = P(E_x \cap E_y)$.

- $P(X = x \text{ and } Y = y)$ is joint probability mass function of $X$ and $Y$

Conditional probability:

- $P(X = x|Y = y) = \frac{P(X = x \text{ and } Y = y)}{P(Y = y)} = \frac{P(E_x \cap E_y)}{P(E_y)} = P(E_x|E_y)$
Example

Roll a die (again):

- Sample space $S = \{1, 2, 3, 4, 5, 6\}$
- $E = \{1, 2\}$ is event that a 1 or a 2 is observed. $F = \{2, 3\}$ that a 2 or 3 is observed.
- Random variable $X = 1$ on event $E$ and 0 otherwise. Random variable $Y = 1$ on event $F$ and 0 otherwise.
- $P(X = 1 \text{ and } Y = 1) = P(E \cap F) = P(\{2\}) = \frac{1}{6}$
- $P(X = 1 \text{ and } Y = 0) = P(E \cap F^c) = P(\{1, 2\} \cap \{1, 4, 5, 6\}) = P(\{1\}) = \frac{1}{6}$
- $P(X = 0 \text{ and } Y = 1) = P(E^c \cap F) = P(\{3, 4, 5, 6\} \cap \{2, 3\}) = P(\{3\}) = \frac{1}{6}$
- $P(X = 0 \text{ and } Y = 0) = P(E^c \cap F^c) = P(\{4, 5, 6\}) = \frac{3}{6}$
- $P(Y = 0) = P(F^c) = P(\{1, 4, 5, 6\}) = \frac{4}{6}$. $P(Y = 1) = \frac{2}{6}$
- $P(X = 0|Y = 0) = \frac{P(X = 0 \text{ and } Y = 0)}{P(Y = 0)} = \frac{\frac{3}{6}}{\frac{4}{6}} = \frac{3}{4}$. $P(X = 1|Y = 0) = \frac{1}{4}$
Chain rule: $P(X = x \text{ and } Y = y) = P(X = x|Y = y)P(Y = y)$.

Consequences of chain rule:

- **Marginalisation:**
  Suppose RV $Y$ takes values in $\{y_1, y_2, \ldots, y_n\}$. Then
  
  $$P(X = x) = P(X = x \text{ and } Y = y_1) + \cdots + P(X = x \text{ and } Y = y_n)$$
  
  $$= \sum_{i=1}^{n} P(X = x|Y = y_i)P(Y = y_i)$$

Example: roll a die, $X$ and $Y$ as before.

- $P(X = 0) = P(X = 0|Y = 0)P(Y = 0) + P(X = 0|Y = 1)P(Y = 1) = \frac{3}{4} \times \frac{4}{6} + \frac{1}{2} \times \frac{2}{6} = \frac{2}{3}$

- Double check: $P(X = 0) = P(\{3, 4, 5, 6\}) = \frac{4}{6} = \frac{2}{3}$. 

Review: Random Variables
Consequences of chain rule (cont):

- Bayes rule: \( P(X = x \mid Y = y) = \frac{P(Y = y \mid X = x)P(X = x)}{P(Y = y)} \).
- \( P(X = x) = \sum_{i=1}^{m} P(X = x \text{ and } Y = y_i) \) when RV \( Y \) takes values in \( \{y_1, y_1, \cdots, y_m\} \)

Independence:

- Discrete random variables \( X \) and \( Y \) are independent if \( P(X = x \text{ and } Y = y) = P(X = x)P(Y = y) \) for all \( x \) and \( y \)

Example: roll a die, \( X \) and \( Y \) as before.

- Are \( X \) and \( Y \) independent? Both depend on outcome 2, so not independent. Check: \( P(X = 0 \text{ and } Y = 0) = \frac{3}{6} = 0.5 \) and \( P(X = 0)P(Y = 0) = \frac{4}{6} \times \frac{4}{6} \approx 0.444 \)
Expected Value

The Expected Value of discrete random variable $X$ taking values in \{\(x_1, x_2, \cdots, x_n\}\) is:

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

- **For independent** random variables $X$ and $Y$ then $E[XY] = E[X]E[Y]$.

Example: roll a die, $X$ and $Y$ as before.

- $E[X] = 1 \times P(X = 1) + 0 \times P(X = 0) = 1 \times \frac{2}{6} + 0 \times \frac{4}{6} = \frac{2}{6}$
- $E[3X + 1] = (3 \times 1 + 1)P(X = 1) + (3 \times 0 + 1)P(X = 0) = 4 \times \frac{2}{6} + 1 \times \frac{4}{6} = \frac{12}{6} = 2$
- Double check: $3E[X] + 1 = 3 \times \frac{2}{6} + 1 = 2$
Expected Value

Conditional expectation of $X$ given $Y = y$ is:

$$E[X|Y = y] = \sum_x xP(X = x|Y = y)$$

- **Linearity**: $E[\sum_i Y_i|X = x] = \sum_i E[Y_i|X = x]$
- $E[X] = \sum_y E[X|Y = y]P(Y = y)$

Example: roll a die, $X$ and $Y$ as before.

- $E[X|Y = 0] = 1 \times P(X = 1|Y = 0) + 0 \times P(X = 0|Y = 0) = 1 \times \frac{1}{4} + 0 \times \frac{3}{4} = \frac{1}{4}$
- $E[X|Y = 1] = 1 \times P(X = 1|Y = 1) + 0 \times P(X = 0|Y = 1) = 1 \times \frac{1}{6} + 0 \times \frac{1}{6} = \frac{1}{2}$
- $E[X] = E[X|Y = 0]P(Y = 0) + E[X|Y = 1]P(Y = 1) = \frac{1}{4} \times \frac{4}{6} + \frac{1}{2} \times \frac{2}{6} = \frac{2}{6}$
Variance

The variance of $X$ taking values in $D = \{x_1, x_2, \cdots, x_n\}$ is:

$$Var(X) = \sum_{i=1}^{n} (x_i - \mu)^2 p(x_i) = E[X^2] - (E[X])^2$$

with $\mu = E[X] = \sum_{i=1}^{n} x_i p(x_i)$

- $Var(X) \geq 0$
- Standard deviation is square root of variance $\sqrt{Var(X)}$.
- $Var(aX + b) = a^2 Var(X)$
- For independent random variables $X$ and $Y$ then $Var(X + Y) = Var(X) + Var(Y)$
Covariance

The covariance of $X$ and $Y$ is:

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = E[XY] - E[X]E[Y]$$

- $\text{Cov}(X, X) = \text{Var}(X)$.
- When $X$ and $Y$ are independent then $E[XY] = E[X]E[Y]$ and $\text{Cov}(X, Y) = 0$.

The correlation between $X$ and $Y$ is:

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

- Takes values between -1 and 1.
Summary Statistics

Expected value, variance, covariance and correlation are all examples of summary statistics.

- Expected value $E[X]$ indicates the overall outcome of many repetitions of an experiment (it's a sort of prediction).
- Variance $Var(X)$ indicates the spread of $X$.
- Covariance $Cov(X, Y)$ is positive if $X$ and $Y$ tend to increase together, and negative if an increase in one tends to correspond to a decrease in the other.
- Correlation $Corr(X, Y)$ indicates the strength of a linear relationship between $X$ and $Y$. 
Bernoulli Random Variable

Suppose an experiment results in Success or Failure.

- $X$ is a random indicator variable, $X = 1$ on success, $X = 0$ on failure
- $P(X = 1) = p$
- $P(X = 0) = 1 - p$
- $X$ is a Bernoulli random variable.
- Sometimes write $X \sim Ber(p)$.

Examples:
- Coin flip
- Random binary digit
- Packet erasure in a wireless network
**Binomial Random Variable**

Consider $n$ independent trials of a $Ber(p)$ random variable

- $X$ is the number of successes in $n$ trials
- $X$ is the sum of $n$ Bernoulli random variables,
  \[ X = X_1 + X_2 + \cdots + X_n, \]
  where random variable $X_i \sim Ber(p)$ is 1 if success in trial $i$ and 0 otherwise.
- $X$ is a **Binomial** random variable: $X \sim Bin(n, p)$

$$P(X = i) = \binom{n}{i} p^i (1 - p)^{n-i}, \quad i = 0, 1, \ldots, n$$

(recall $\binom{n}{i}$ is the number of outcomes with exactly $i$ successes and $n - i$ failures)

- $E[X] = np$, $Var(X) = np(1 - p)$ ($X$ is the sum of $n$ independent Bernoulli RVs)

Examples:
- number of heads in $n$ coin flips
- number of 1’s in randomly generated bit string of length $n$
- number of packets erased out of a file of $n$ packets
Example

A shopper wants to compare bottles of wine. From a shelf with 6 bottles labelled A-F, each different, he selects 3 independently and uniformly at random. What is the probability that he picks bottle B?

- Let indicator random variable $X = 1$ if pick bottle B and 0 otherwise. For $X = 1$ there are three cases to consider:
  - Picks B first. Happens with probability $\frac{1}{6}$.
  - Does not pick B first but picks B second. Happens with probability $(1 - \frac{1}{6}) \frac{1}{5}$.
  - Does not pick B first or second but picks B third. Happens with probability $(1 - \frac{1}{6})(1 - \frac{1}{5}) \frac{1}{4}$.
- So $P(X = 1) = \frac{1}{6} + (1 - \frac{1}{6}) \frac{1}{5} + (1 - \frac{1}{6})(1 - \frac{1}{5}) \frac{1}{4} = 0.5$
Example

Joe Lucky plays the lottery on any given week with probability $p$, independently of other weeks. Each time he plays he has probability $q$ of winning. During a period of $n$ weeks, let $X$ be the number of times that he played the lottery and $Y$ the number of times that he won.

- What is the probability that he played the lottery in a week given that he did not win anything that week?
- Let $E$ be the event that he played and $F$ the event that he did not win. Use Bayes Rule.

$$P(E|F) = \frac{P(F|E)P(E)}{P(F)} = \frac{P(F|E)P(E)}{P(F|E)P(E) + P(F|E^c)P(E^c)}$$

$$= \frac{(1 - q)p}{(1 - q)p + 1 \times (1 - p)} = \frac{p - pq}{1 - pq}$$
Example (cont)

Recall $X$ is the number of times that he played the lottery and $Y$ the number of times that he won.

- What is the conditional PMF $P(Y = y | X = x)$?
- Its Binomial:

$$P(Y = y | X = x) = \begin{cases} \binom{x}{y} q^y (1 - q)^{x-y} & 0 \leq y \leq x \\ 0 & \text{otherwise} \end{cases}$$
Example

Since there is no direct flight from Dublin (D) to Atlanta (A) you need to travel via either Chicago (C) or New York (N). Flights from D to C and from C to A independently are delayed by 2 hours with probability $p$. Flights from D to N and from N to A independently are delayed by 1 hour with probability $q$. Each time you fly you choose to fly via C or N with equal probability.

• What is the average delay from D to A ?
• Let random variable $X_{DC}$ be the delay D to C (either 0 or 2), similarly $X_{CA}$, $X_{DN}$ and $X_{NA}$.

\[
E[\text{delay}] = E[\text{delay}|C]P(C) + E[\text{delay}|N]P(N)
= E[X_{DC} + X_{CA}]P(C) + E[X_{DN} + X_{NA}]P(N)
= E[X_{DC}]P(C) + E[X_{CA}]P(C) + E[X_{DN}]P(N) + E[X_{NA}]P(N)
= 2p \cdot \frac{1}{2} + 2p \cdot \frac{1}{2} + q \cdot \frac{1}{2} + q \cdot \frac{1}{2} = 2p + q
\]
Example (cont)

Suppose you arrive with delay 2 hours. What is the probability that you travelled via New York?

- Let $E$ the event that travelled via New York and $F$ be the event that delayed 2 hours. Use Bayes:

$$P(E|F) = \frac{P(F|E)P(E)}{P(F)} = \frac{P(F|E)P(E)}{P(F|E)P(E) + P(F|E^c)P(E^c)}$$

$$= \frac{q^{2 \frac{1}{2}}}{q^{2 \frac{1}{2}} + 2p(1-p)^{\frac{1}{2}}} = \frac{q^2}{q^2 + 2p(1-p)}$$