Today: Independent trials

Next: ASV 3.3

Week 4:

- homework 3 (due Wednesday October 26)
**Independent random variables** \[ P(AB) = P(A)P(B) \]

A collection \( X_1, X_2, \ldots, X_n \) of random variables defined on the same sample space are independent if for any \( B_1, B_2, \ldots, B_n \subseteq \mathbb{R} \), the events 
\[
\{X_1 \in B_1\}, \{X_2 \in B_2\}, \ldots, \{X_n \in B_n\} \text{ are independent}
\]
i.e., 
\[
P(\{X_1 \in B_1\} \cap \{X_2 \in B_2\} \cap \cdots \cap \{X_n \in B_n\}) = P(X_1 \in B_1) \cdot P(X_2 \in B_2) \cdots P(X_n \in B_n)
\]

**Special case**: if \( X_j \) are discrete random variables, it suffices to check the simpler condition for any real numbers \( t_1, t_2, \ldots, t_n \)
\[
P(X_1 = t_1, X_2 = t_2, \ldots, X_n = t_n) = P(X_1 = t_1) \cdot P(X_2 = t_2) \cdots P(X_n = t_n)
\]

**Example**: Let \( X_1, X_2, \ldots, X_n \) be fair coin tosses, \( H \sim 1, T \sim 0 \)
\[
P(X_1 = t_1, X_2 = t_2, \ldots, X_n = t_n) = \frac{1}{2^n} = \frac{1}{2} \cdot \frac{1}{2} \cdots \frac{1}{2} = P(X_1 = t_1) \cdots P(X_n = t_n)
\]
Experiments can have numerical observables, but sometimes you only observe whether there is success or failure. We model this with a random variable $X$ taking value 1 with probability $p$, and value 0 with probability $1-p$.

$$X \sim \text{Ber}(p) \quad (\text{Bernoulli})$$

In practice, we usually repeat the experiment many times, making sure to use the same setup each trial. The previous trials do not influence the future ones.
Independent trials. Binomial distribution $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$

Let $X_1, X_2, \ldots, X_n$ be independent $\text{Ber}(p)$ random variables.

E.g. $P(X_1=0, X_2=1, X_3=1, X_4=0, X_5=0, X_6=0)$

$= P(X_1=0) \cdot P(X_2=1) \cdot P(X_3=1) \cdot P(X_4=0) \cdot P(X_5=0) \cdot P(X_6=0)$

$= (1-p) \cdot p \cdot p \cdot (1-p) \cdot (1-p) \cdot (1-p) = p^n (1-p)^5$

Run $n$ independent trials, each with success probability $p$, $X_1, X_2, \ldots, X_n \sim \text{Ber}(p)$

Let $S_n = \# \text{ successful trials} = X_1 + X_2 + \ldots + X_n$

What is the distribution of $S_n$?

$P(S_n = k) = P( \{ \text{exactly } k \text{ of the } n \text{ trials are successful}\} )$

$= \binom{n}{k} p^k (1-p)^{n-k}$ Binomial distribution $\text{Bin}(n, p)$

If $p = \frac{1}{2}, (1-p) = \frac{1}{2}$, $P(S_n=k) = \frac{\binom{n}{k}}{2^n}$, $\# \text{ of Heads in } n \text{ tosses} \sim \text{Bin}(n, \frac{1}{2})$
Independent trials. Binomial distribution

**Example** Roll a fair die 10 times. What is the probability that 6 comes up at least 3 times?

- Success: $X_1, X_2, \ldots, X_{10} \sim \text{Ber} \left( \frac{1}{6} \right)$
- $S_{10} = X_1 + X_2 + \cdots + X_{10} \sim \text{Bin} \left( 10, \frac{1}{6} \right)$

\[
P(S_{10} \geq 3) = \sum_{k=3}^{10} P(S_{10} = k) = 1 - P(S_{10} = 0) - P(S_{10} = 1) - P(S_{10} = 2)
\]

\[
= 1 - \binom{10}{0} \left( \frac{1}{6} \right)^0 \left( \frac{5}{6} \right)^{10} - \binom{10}{1} \left( \frac{1}{6} \right)^1 \left( \frac{5}{6} \right)^9 - \binom{10}{2} \left( \frac{1}{6} \right)^2 \left( \frac{5}{6} \right)^8 
\approx 22.5\% 
\]

What is the probability that no 6 is rolled in the 10 rolls?

\[
P(S_{10} = 0) = \binom{10}{0} \left( \frac{1}{6} \right)^0 \left( \frac{5}{6} \right)^{10} = \left( \frac{5}{6} \right)^{10}
\]
First success time. Geometric distribution

Keep rolling. Let $N$ denote the first roll where a 6 appears. $N$ is a random variable.

What is the distribution of $N$?

$N =$ first success in repeated independent trials (success rate $p$).

Model trials with (unlimited number of) independent Ber($p$)'s $X_1, X_2, X_3, \ldots$, $N \in \{1, 2, 3, 4, \ldots \}$

$\{N = k\} = \{X_1 = 0, X_2 = 0, \ldots, X_{k-1} = 0, X_k = 1\}$

$$P(N=k) = P(X_1 = 0) \cdot P(X_2 = 0) \cdot \cdots \cdot P(X_{k-1} = 0) \cdot P(X_k = 1)$$

$$= (1-p)(1-p) \cdots (1-p) \cdot p = (1-p)^{k-1} \cdot p$$

Geometric Distribution $\text{Geom}(p)$ on $\{1, 2, 3, \ldots \}$, is it?

$$\sum_{k=1}^{\infty} P(N=k) = \sum_{k=1}^{\infty} (1-p)^{k-1} \cdot p = P \cdot \sum_{k=0}^{\infty} (1-p)^k = P \cdot \frac{1}{1-(1-p)} = p \cdot \frac{1}{p} = 1$$
Rare events. Poisson distribution

Let \( \lambda > 0 \) and let \( X \) be a r.v. taking values in \( \{0, 1, 2, \ldots\} \). \( X \) has Poisson distribution with parameter \( \lambda \) if

\[
P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{for} \quad k \in \{0, 1, 2, \ldots\}
\]

We write \( X \sim \text{Pois}(\lambda) \).

Poisson distribution describes the probability that a "rare" event occurs \( k \) times after repeating the experiment (independent trials) "many" times.

Is this a probability distribution? Yes

\[
e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}
\]

\( P(X = k) \geq 0 \),

\( \lambda \) gives the "expected number" of occurrences.