Probability Formulas

Conditional Prob.:	$\mathbb{P}(A\|B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$,	Law Of Total Prob.:	$\mathbb{P}(A) = \sum_{i \geq 1} \mathbb{P}(A\|B_i)\mathbb{P}(B_i)$,
Baye's Theorem:	$\mathbb{P}(A\|B) = \mathbb{P}(B\|A)\dfrac{\mathbb{P}(A)}{\mathbb{P}(B)}$.	Stirling's Formula:	$n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$.
Inclusion-Exclusion:	$\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B)$.	Expectation:	$\mathbb{E}[X] = \iint x f_{X,Y}(x,y)\, dx\, dy$

Independence:

Two variables X and Y are *independent* iff:

- $\mathbb{P}(A|B) = \mathbb{P}(A)\mathbb{P}(B)$.
- $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.
- $p_{X,Y}(x,y) = p_X(x)p_Y(y)$ or $f_{X,Y}(x,y) = f_X(x)f_Y(y)$.

The Multiplication Rule:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A)$$

or in general, for random variables $A_1, A_2, \ldots, A_n$:

$$\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\ldots\mathbb{P}(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

Law of the Unconscious Statistician, (AKA "LOTUS"):

$$\mathbb{E}[g(X)] = \sum_x g(x)f_X(x), \quad \text{or, for continuous variables:} \quad \mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x)\, dx.$$

Joint Distributions:

$$p_X(x) = \sum_y p_{X,Y}(x,y) \quad \text{and} \quad p_Y(y) = \sum_x p_{X,Y}(x,y).$$

Conditional Distributions:

$$p_{X|\{Y=y\}}(x) = \frac{p_{X,Y}(x,y)}{p_Y(y)}, \qquad \mathbb{E}\big[X|Y=y\big] = \sum_x x p_{X|\{Y=y\}}(x),$$

$$p_{X|B}(x) = \mathbb{P}(X=x|B) = \frac{\mathbb{P}(\{X=x\} \cap B)}{\mathbb{P}(B)}.$$

Covariance and Correlation:

$$\text{cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y], \quad \text{and} \quad cor(X,Y) = \frac{\text{cov}(X,Y)}{\sqrt{Var(X)Var(Y)}}.$$

NOTE: X and Y are independent $\Rightarrow \text{cov}(X,Y) = 0$ is a **one way implication**.

$$Var(X+Y) = Var(X) + Var(Y) + 2\text{cov}(X,Y).$$

Generating Functions:

For two *independent* random variables X and Y:

$$\mathbb{E}\left[s^X\right] = G_X(s) = \sum_{k=0}^{n} s^k \mathbb{P}(X=k), \qquad \mathbb{E}[X] = G'_X(1),$$

$$G_{X+Y}(s) = G_X(s)G_Y(s). \qquad Var(X) = G''_X(1) - G'_X(1) + (G'_X(1))^2.$$

NOTE: If you roll a die n times, then the probability of getting a total of exactly k is the coefficient of s^k term in the generating function.

Chebechev's Inequality:

$$\mathbb{P}(|X-\mu| \geq a) \leq \frac{Var(X)}{a^2}.$$

NOTE: You can use Chebyshev's inequality to to find a value of n such that if you toss a coin n times the proportion of heads will be within 0.01 of 0.5 with probability at least 0.95. So you'd check: $1-\mathbb{P}(|X-0.5| \geq 0.01) \geq \frac{Var(X)}{(0.01)^2}$.

The Weak Law of Large Numbers:

For some independent random variables $A_n = \underbrace{A+A+\ldots+A}_{n\ times}$, if $a>0$ then:

$$\mathbb{P}(A_n - \mu \geq a) \to 0 \quad \text{as} \quad n \to \infty.$$

TABLE OF DISTRIBUTIONS:

Discrete	P.D.F, f(x)	Expectation	Variance	P.G.F, $G_X(s)$
Bernoulli	$\mathbb{P}(X=x) = \begin{cases} p & \text{if } x=1 \\ 1-p & \text{if } x=0 \end{cases}$	p	$p(1-p)$	$(1-p)+ps$
Binomial, $X \sim B(n,p)$	$\mathbb{P}(X=k) = \binom{n}{k} p^k(1-p)^{n-k}$	np	$np(1-p)$	$((1-p)+ps)^n$
Geometric, $X \sim Geom(p)$	$\mathbb{P}(X=k) = p(1-p)^{k-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{ps}{1-s(1-p)}$
Poisson, $X \sim Po(\lambda)$	$\mathbb{P}(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}$	λ	λ	$e^{\lambda(s-1)}$
Continuous	**P.D.F, f(x)**	**Expectation**	**Variance**	**C.D.F, F(x)**
Uniform, $X \sim U[a,b]$	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{x-a}{b-a}$
Normal, $X \sim N(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}$	μ	σ^2	$\Phi\left(\frac{x-\mu}{\sigma}\right)$
Standard Normal, $X \sim N(0,1)$	$\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$	0	1	$\Phi(X)$
Exponential, $X \sim exp(\lambda)$	$\lambda e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$1-e^{-\lambda x}$

Advice for if You're Stuck

- Try using the **multiplication rule**:
$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(B|A)\mathbb{P}(A)$$
- If the events are disjoint, use
$$\mathbb{P}(A_1 \cup A_2 \cup \cdots \cup A_n) = \mathbb{P}(A_1) + \mathbb{P}(A_2) \ldots \mathbb{P}(A_n).$$
- Try taking **complements**:
$$\mathbb{P}(A_1 \cup A_2 \cup \cdots \cup A_n) = 1 - \mathbb{P}((A_1 \cup A_2 \cup \cdots \cup A_n)^c) = 1 - \mathbb{P}(A_1^c \cap A_2^c \cap \cdots \cap A_n^c).$$
- Try using the **Inclusion-Exclusion Principle**:
$$\mathbb{P}(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n}(\mathbb{P}(A_i)) + \sum_{i<j}^{n}(\mathbb{P}(A_i \cap A_j)) + \cdots + (-1)^{n+1}\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n).$$
- If you can't calculate $\mathbb{P}(A)$ directly, try splitting it up and using the **Law of Total Probability.**

Questions:

Question, Bounds for $\mathbb{P}(A|B)$:

If $\mathbb{P}(A) = \frac{1}{2}$ and $\mathbb{P}(B) = \frac{4}{5}$ then find the upper and lower bounds of $\mathbb{P}(A|B)$.

Solution: We have two cases to check. Either:
$A \subseteq B$**:** Which gives that $\mathbb{P}(A \cap B) = \mathbb{P}(A)$ which means that $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{5}{8}$.
Or we have:
$\Omega = A \cup B$**:** Hence $\mathbb{P}(A \cup B) = \mathbb{P}(\Omega) = 1$ and so $\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B) = \frac{1}{2} + \frac{4}{5} - 1 = 0.3$
and hence $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{0.3}{0.8} = \frac{3}{8}$.
Finally then, we have
$$\frac{3}{8} \leq \mathbb{P}(A|B) \leq \frac{5}{8}.$$

Question, $\mathbb{E}$ number of consecutive Heads:
What is the expected number of pairs of consecutive heads if you toss a coin n times?

Solution: Let X_j be the random variable such that

$$X_j = \begin{cases} 1 & \text{if the } j\text{-th and } (j+1)\text{-th tosses are both heads} \\ 0 & \text{otherwise} \end{cases}$$

Where $1 \leq j \leq n-1$. For each X_j we have that $\mathbb{E}[X_j] = \frac{1}{4}$ since the chance of two heads is $\mathbb{P}(\text{HH}) = \frac{1}{4}$. We're interested in $\sum_{j=1}^{n-1} X_j$, so

$$\mathbb{E}\left[\sum_{j=1}^{n-1} X_j\right] = \sum_{j=1}^{n-1} \mathbb{E}[X_j] = \frac{n-1}{4}.$$

Question, $\mathbb{E}[XY]$:
I toss a coin 3 times. X represents the number of heads in the *first two tosses*, and Y represents the number of heads in the *last two tosses*. Find $\mathbb{E}[XY]$.

Solution: Let's list all the possible outcomes (*left*) and the corresponding value of XY (*right*):

HHH	THH	TTH	4	2	0
HHT	HTT	TTT	2	0	0
HTH	THT	–	1	1	–

Now we just use the fact that $\mathbb{E}[XY] = \sum xy\mathbb{P}(XY = xy) = 4 \cdot \frac{1}{8} + 2 \cdot \frac{2}{8} + 1 \cdot \frac{2}{8} + 0 \cdot \frac{3}{8} = \frac{10}{8}$.