PROBABILITY FACTS CHEAT SHEET

Fact. *(**Basic Counting Principle**) Suppose 2 experiments are to be performed.*
If one experiement can result in m possibilities
Second experiment can result in n possibilities
Then together there are mn possibilities

Fact. *If $r \leq n$, then*

$$\binom{n}{r} = \frac{n!}{(n-r)!r!}$$

*and say n choose r, represents the number of possible **combinations** of objects taken r at a time.*
($\star$) Order DOES NOT Matter her

- With n objects. There are
$$n(n-1)\cdots 3\cdot 2\cdot 1 = n!$$
different **permutations** of the n objects.
- Binomial Theorem: $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$
- **Probability** will be a rule given by the following Axioms (Laws that we all agree on)
  - Let S be a sample space.
  - A probability will be a function $\mathbb{P}(E)$ where the input is a set/event $E \subset S$ such that
  - **Axiom 1:** $0 \leq \mathbb{P}(E) \leq 1$ for all events E.
  - **Axiom 2:** $\mathbb{P}(S) = 1$.
  - **Axiom 3:** (disjoint property) If the events $E_1, E_2, \ldots$ are pairwise disjoint/mutually exclusive then
$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i).$$
    * Mutually exclusive means that $E_i \cap E_j = \emptyset$ when $i \neq j$.

Proposition 1. *(a)* $\mathbb{P}(\emptyset) = 0$
(b) If $A_1, \ldots, A_n$ are pairwise disjoint, $\mathbb{P}(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mathbb{P}(A_i)$.
(c) $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$.
(d) If $E \subset F$, then $\mathbb{P}(E) \leq \mathbb{P}(F)$.
(e) $\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F)$.

- We say E and F are **independent events** if
$$\mathbb{P}(E \cap F) = \mathbb{P}(E)\mathbb{P}(F).$$
- $\mathbb{P}(E \mid F) = \frac{\mathbb{P}(E\cap F)}{\mathbb{P}(F)}$ and $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B \mid A)$
- **The Law of Total Probability:** If $F_1, \ldots, F_n$ are mutually exclusive events such that they make up everythinn $S = \bigcup_{i=1}^{n} F_i$ then
$$\mathbb{P}(E) = \sum_{i=1}^{n} \mathbb{P}(E \mid F_i)\mathbb{P}(F_i).$$
- **Bayes' formula:** If $S = \bigcup_{i=1}^{n} F_i$, for any any j,
$$\mathbb{P}(F_j \mid E) = \frac{\mathbb{P}(E \mid F_j)\mathbb{P}(F_j)}{\sum_{i=1}^{n} \mathbb{P}(E \mid F_i)\mathbb{P}(F_i)}.$$

  - When $n = 2$, with $S = F_1 \cup F_2$, then
$$\mathbb{P}(F_1 \mid E) = \frac{\mathbb{P}(E \mid F_1)\mathbb{P}(F_1)}{\mathbb{P}(E \mid F_1)\mathbb{P}(F_1) + \mathbb{P}(E \mid F_2)\mathbb{P}(F_2)}$$
and
$$\mathbb{P}(F_2 \mid E) = \frac{\mathbb{P}(E \mid F_2)\mathbb{P}(F_2)}{\mathbb{P}(E \mid F_1)\mathbb{P}(F_1) + \mathbb{P}(E \mid F_2)\mathbb{P}(F_2)}$$

- **Discrete random variable:**
  - **PMF (Probability Mass Function):** $p_X(x) := \mathbb{P}(X = x)$, (NOTE: some texts may use the notation for $f_X(x) = \mathbb{P}(X = x)$ to denote the PMF)
    * **Properties of a pmf $p(x)$:**
    * Note that we must have $0 < p(x_i) \leq 1$ for $,i = 1, 2, \ldots.$ and $p(x) = 0$ for all other values of x can't attain.
    * Also must have
$$\sum_{i=1}^{\infty} p(x_i) = 1.$$
  - **CDF:** $F_X(x) := \mathbb{P}(X \leq x)$.
- **Infinite Series: Geometric Series**
$$1 + x + x^2 + x^3 + \cdots + = \frac{1}{1-x}.$$
then differentiating we have
$$0 + 1 + 2x + 3x^2 + \cdots + = \frac{1}{(1-x)^2}.$$
- **Continuous Random Variables:**

Definition. A random variable X is said to have a **continuous distribution** if there exists a nonnegative function f_X (called the probability distribution function or **PDF**)such that
$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x)dx$$
for every a and b. [Sometimes we write that for nice sets $B \subset \mathbb{R}$ we have $\mathbb{P}(X \in B) = \int_B f_X(x)dx$.]

  - All **PDFs** must satisfy:
(1) $f(x) \geq 0$ for all x
(2) $\int_{-\infty}^{\infty} f(x)dx = 1$.
  - **CDF:** $F_X(x) := \mathbb{P}(X \leq x)$
- **Expected Values:** If $g : \mathbb{R} \to \mathbb{R}$
  - Discrete R.V.: List $X \in \{x_1, x_2, \ldots\}$
    * $\mathbb{E}[g(X)] = \sum_{i=1}^{\infty} g(x_i)\, p_X(x_i)$
  - Continuous R.V.:
    * $\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)\, f_X(x)\, dx$
- **Fact:** For continuous R.V we have the following useful relationship
  - Since $F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f_X(y)dy$ then by the fundamental theorem of calculus we have
$$F_X'(x) = f_X(x).$$
- How to find the PDF of $Y = g(X)$ where X is the PDF of X.
  - **Step1:** First start by writing the cdf of Y and in terms of F_X:
  - **Step2:** Then use the relation $f_Y(y) = F_Y'(y)$ and take a derivative of the expression obtained in Step 1.
- **Joint Distributions:**
  - **Discrete: joint probability mass(discrete density) function**
$$p(x, y) = \mathbb{P}(X = x, Y = y).$$
    * Some texts may use $f(x, y)$ to denote the PMF.

- **Continuous:**For random variables X, Y we let $f(x, y)$ be the **joint probability density function,** if
$$\mathbb{P}(a \le X \le b, c \le Y \le d) = \int_a^b \int_c^d f(x,y)dydx.$$
  - Or in general if $D \subset \mathbb{R}^2$ is a region in the plane then
$$\mathbb{P}(X \in D) = \int\int_D f(x,y)dydx.$$
- We also have the **multivariate cdf:($\star\star$)** defined by
$$F_{X,Y}(x,y) = \mathbb{P}(X \le x, Y \le y).$$
- **INDEPENDENCE:**
- Discrete: We say discrete R.V. X, Y are **independent** if
$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y),$$
for every x, y in the range of X and Y.
  - This is the same as saying that X, Y ar independent if the joint pmf splits into the marginal pmfs: $p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y)$
- Continuous: We say continuous r.v. X, Y are independent if
$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$
for any set A, B

Theorem 2. *Continuous (discrete) r.v. X, Y are independent if and only if their joint pdf (pmf) can be expressed as*
$$f_{X,Y}(x,y) = f_X(x)f_Y(y). \text{ (Continuous Case)},$$
$$p_{X,Y}(x,y) = p_X(x)p_Y(y) \text{ (Discrete Case)}.$$

- **Joint Expectations:** Let $g : \mathbb{R}^2 \to \mathbb{R}$ then
$$\mathbb{E}[g(X,Y)] = \sum_y \sum_x g(x,y)p(x,y), \quad \text{(discrete)}$$
$$\mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy, \text{ Continuous}$$
- Let $g : \mathbb{R}^2 \to \mathbb{R}$. Let X be a discrete random variable and Y be a continuous random variable. If X has pmf $p_X(x)$,Y has joint pdf $f_Y(y)$, and X, Y are independent then
$$\mathbb{E}[g(X,Y)] = \sum_{i=1}^{\infty}\int_{-\infty}^{\infty} g(x_i,y)p_X(x)f_Y(y)dy,$$
- If X, Y are independent and $g, h : \mathbb{R} \to \mathbb{R}$ then
$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}[g(X)]\mathbb{E}[h(Y)].$$

- The **covariance** between X and Y, is defined by
$$\text{Cov}(X,Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)],$$
$$\text{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}X\mathbb{E}Y$$
- For each random variable X, we can define its **moment generating function** $m_X(t)$ by
$$m_X(t) = \mathbb{E}\left[e^{tX}\right] = \begin{cases} \sum_{x_i} e^{tx_i}p(x_i) & \text{,if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} e^{tx}f(x)ds & \text{,if } X \text{ is continuous} \end{cases}.$$
  - Fact 1: If $m_X(t) = m_Y(t) < \infty$ for all t in an interval, then X and Y have the same distribution.
  - Fact 2: If X, Y are independent then $m_{X+Y}(t) = m_X(t)m_Y(t)$.

  - Fact 3: If $m_X(t)$ is the MDF of X then the nth moment of X can be found by
$$\mathbb{E}\left[X^n\right] = m_X^{(n)}\left(0\right).$$

- **How we use CLT (CENTRAL LIMIT THEOREM):** That is, for any random variable X with $\mathbb{E}X = \mu$ then standard deviation σ then
$$\mathbb{P}\left(\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \leq b\right) \approx \mathbb{P}\left(Z \leq b\right) = \Phi(b).$$