The average value of $X$

- Also known as
  - the **expectation**
  - the mean
  - the first moment

- $E(X) = \int x \, dF(x) = \begin{cases} \sum x \, f(x) & \text{if } X \text{ is discrete} \\ \int x \, f(x) \, dx & \text{if } X \text{ is continuous} \end{cases}$

- The $k^{th}$ moment of $X$ is $E(X^k)$ assuming $E(|X|^k) < \infty$

If $Y = r(X)$, then $E(Y) = E(r(X)) = \text{replace } x \text{ with } r(x) \text{ in summation/integration.}$
Properties of expectations

- If $X_1, \ldots, X_n$ are random variables and $a_1, \ldots, a_n$ are constants, then
  \[ E\left( \sum_i a_i X_i \right) = \sum_i a_i E(X_i) \]

- If $X_1, \ldots, X_n$ are independent random variables, then
  \[ E\left( \prod_i X_i \right) = \prod_i E(X_i) \]
The spread of $X$

- Commonly measured using
  - the variance
- If $X$ has mean $\mu$, the variance of $X$ is
  $$V(X) = \sigma^2 = \mathbb{E}((X - \mu)^2) = \int (x - \mu)^2 dF(x)$$

- A well defined variance has the following properties:
  \[
  V(X) = \mathbb{E}(X^2) - \mu^2 \\
  V(aX + b) = a^2 \, V(X) \\
  V \left( \sum_i a_i X_i \right) = \sum_i a_i^2 \, V(X_i)
  \]

standard deviation $sd(X) = \sigma$
Sample statistics

- The sample mean is \( \bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \)

- The sample variance is \( S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 \)

- Let \( X_1, \ldots, X_n \) be independent identically distributed, then
  \[
  \begin{align*}
  E(\bar{X}_n) &= E(X_i) = \mu \\
  V(\bar{X}_n) &= V(X_i) = \frac{\sigma^2}{n} \\
  E(S_n^2) &= \sigma^2
  \end{align*}
  \]

E.g. Let \( X \sim \text{binomial}(n,p) \). We write \( X = \sum(X[i], i=1..n) \), where \( X[i] = 1 \) if toss \( i \) is heads and \( X[i] = 0 \) otherwise. Then \( X = \sum(X[i], i=1..n) \). Also, \( P(X[i]=1) = p \) and \( P(X[i]=0) = 1-p \). Since \( E(X[i]) = (p*1) + ((1-p)*0) = p \), we have \( E(X[i]^2) = (p*1^2) + ((1-p)*0^2) = p \).
The covariance and correlation

- The covariance
  \[ \text{Cov}(X, Y) = \mathbb{E}\left[ (X - \mu_X)(Y - \mu_Y) \right] \]
  \[ = \mathbb{E}(X Y) - \mathbb{E}(X)\mathbb{E}(Y) \]

- The correlation
  \[ \rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}, \quad \text{where} \ |\rho(X, Y)| < 1 \]

- \[ \mathbb{V}(\sum_i a_i X_i) = \sum_i a_i^2 \mathbb{V}(X_i) + 2 \sum_{i < j} a_i a_j \text{Cov}(X_i, X_j) \]
Conditional expectation

- $E(X \mid Y = y) = \begin{cases} \sum x f_{X \mid Y}(x \mid y) & \text{if discrete} \\ \int x f_{X \mid Y}(x \mid y) \, dx & \text{if continuous} \end{cases}$

- $E(E(Y \mid X)) = E(Y)$

- The conditional variance is
  
  $V(Y \mid X = x) = \int (y - \mu(x))^2 f(y \mid x) \, dy$

  $V(Y) = E(V(Y \mid X)) = V(E(Y \mid X))$

$E(r(X,Y) \mid Y=y) = E(r(X)) = \text{replace } x \text{ with } r(x,y) \text{ in summation/integration.}$
Laplace transform of $X$

- Also known as
  - the moment generating function (MGF)
- $\psi_X(t) = E(e^{itX}) = \int e^{itX} dF(X)$

- Properties:
  If $Y = aX + b$, then $\psi_Y(t) = e^{itb} \psi_X(at)$
  If $X_1, \ldots, X_n$ are independent and $Y = \sum X_i$, then $\psi_Y(t) = \prod_i \psi_i(t)$