Random variables and probability mass/density functions

r.v.

\[ X : \mathbb{S} \rightarrow \mathbb{R} \]

probability mass function (density)

\[ p_X(x) = \Pr(X = x) \]

discrete

\[ X = \begin{cases} 1 & \frac{3}{4} \\ 0 & \frac{1}{4} \end{cases} \]

\[ p_X(1) = \frac{3}{4} \]

\[ p_X(2) = 0 \]
binomial random variable

PMF for $X \sim \text{Bin}(100, 0.5)$

$X$ counts the number of heads in $n$ independent coin tosses, each with probability $p$ of coming up heads.

$E(X) = \sum_{k=0}^{100} k \cdot p_X(k)$

$p.m.f.$ for $\text{Bin}(100, 0.5)$

$k = 53$
binomial tails

For a random variable $X$, the tails of $X$ are the parts of the PMF that are “far” from its mean.

PMF for $X \sim \text{Bin}(100,0.5)$
For a random variable $X$, the tails of $X$ are the parts of the PMF that are “far” from its mean.
Often, we want to bound the probability that a random variable $X$ is “extreme.”

Such a bound is called a “tail bound”.
Suppose we know that \( X \) is always non-negative.

**Theorem:** If \( X \) is a non-negative random variable, then for every \( \alpha > 0 \), we have

\[
P(X \geq \alpha) \leq \frac{E[X]}{\alpha}
\]

Equivalently

\[
P(X \geq cE[X]) \leq \frac{1}{c}
\]
Markov’s inequality

**Theorem:** If $X$ is a non-negative random variable, then for every $\alpha > 0$, we have

$$P(X \geq \alpha) \leq \frac{E[X]}{\alpha}$$

**Proof:**

$$E[X] = \sum x \times p(x)$$

$$= \sum_{x<\alpha} x \times p(x) + \sum_{x\geq\alpha} x \times p(x)$$

$$\geq 0 + \sum_{x\geq\alpha} \alpha p(x) \quad (x \geq 0; \alpha \leq x)$$

$$= \alpha P(X \geq \alpha)$$
Variance and Chebyshev’s inequality

If we know more about a random variable, we can often use that to get better tail bounds.

Suppose we also know the variance.

\[ \text{Var}[Y] = E[(Y - E(Y))^2] \]

Standard deviation = square root of variance
\[ \text{Var}[Y] = E[(Y - E(Y))^2] = E[(Y - \mu)^2] \]

PMF for \( X \sim \text{Bin}(100, 0.5) \)
Chebyshev’s inequality

If we know more about a random variable, we can often use that to get better tail bounds.

Suppose we also know the variance.

**Theorem:** If $Y$ is an arbitrary random variable with $E[Y] = \mu$, then, for any $\alpha > 0$,

$$P(|Y - \mu| \geq \alpha) \leq \frac{\text{Var}[Y]}{\alpha^2}$$

$$\alpha = \sigma$$

$$\leq \frac{\sigma^2}{(\sigma^2)^2} = \frac{1}{\sigma^2}$$

$$\Pr(|Y - \mu| \geq \sigma) \leq \frac{1}{\sigma^2}$$
**Theorem:** If $Y$ is an arbitrary random variable with $\mu = E[Y]$, then, for any $\alpha > 0$,

$$P(|Y - \mu| \geq \alpha) \leq \frac{\text{Var}[Y]}{\alpha^2}$$

**Proof:** Let $X = (Y - \mu)^2$

$X$ is non-negative, so we can apply Markov’s inequality:

$$P(|Y - \mu| \geq \alpha) = P(X \geq \alpha^2)$$

$$\leq \frac{E[X]}{\alpha^2} = \frac{\text{Var}[Y]}{\alpha^2}$$
Important Facts about Variance

**Linearity of expectation** always holds, i.e.,

\[ E(aX_1 + a_2X_2 + \ldots + a_nX_n) = a_1E(X_1) + a_2E(X_2) + \ldots + a_nE(X_n) \]

**Linearity of variance** holds **only** if the random variables are independent.

\[ \text{Var}[X_1 + \ldots + X_n] = \text{Var}[X_1] + \ldots + \text{Var}[X_n] \]

Also, if \( a \) is a constant, then

\[ \text{Var}[a \cdot X] = a^2 \cdot \text{Var}[X] \]
Performance/estimation using repetition

\[ \bar{x}_n = \frac{1}{n} \left( x_1 + x_2 + \ldots + x_n \right) \]

\[ E(\bar{x}_n) = \frac{1}{n} \sum_{i=1}^{n} E(x_i) \]

\[ = \frac{1}{n} \cdot n \cdot p = p \]

\[ \text{Var}(\bar{x}_n) = \text{Var}\left( \frac{1}{n} \sum_{i=1}^{n} x_i \right) \]

\[ = \frac{1}{n^2} \text{Var}\left( \sum_{i=1}^{n} x_i \right) \]

\[ \text{indep} \]

\[ \text{Var}(\bar{x}_n) = \frac{n \cdot \sigma^2}{n^2} = \frac{\sigma^2}{n} \]

\[ \text{Law of large numbers} \]

\[ \sum \text{ independent with mean } \mu \]

\[ \frac{1}{n} \sum_{i=1}^{n} x_i \rightarrow E(x_i) \]

\[ \text{Pr}\left( \left| \bar{x}_n - \mu \right| \geq c \right) \leq \frac{\text{Var}(\bar{x}_n)}{c^2} \]

\[ \frac{\sigma^2}{n} \rightarrow 0 \]