Useful facts

Mathematical preliminaries

1. Arithmetic, geometric, and harmonic series:

\[
1 + 2 + \cdots + n = \frac{n(n+1)}{2}
\]

\[
1 + r + r^2 + \cdots = \frac{1}{1-r} \quad \text{if } 0 < r < 1
\]

\[
1 + \frac{1}{2} + \cdots + \frac{1}{k} \approx \ln k
\]

2. \(1 + x \leq e^x\) and \(1 - x \leq e^{-x}\). When \(x\) is small, this is a good approximation.

Multiple events, conditioning, and independence

1. Events \(A\) and \(B\) are independent if \(\Pr(A \cap B) = \Pr(A)\Pr(B)\).

2. Union bound: \(\Pr(A_1 \cup \cdots \cup A_k) \leq \Pr(A_1) + \cdots + \Pr(A_k)\).

3. Suppose you throw \(m\) balls into \(n\) bins.
   - (Birthday paradox) If \(m \geq \sqrt{n}\), then there will very likely be a bin with more than one ball.
   - (Coupon collector) If \(m \geq n \log n\), then very likely there will be no empty bins.
   - If \(m = n\), then the largest bin will contain roughly \(\log n\) balls.

4. Conditional probability formulas
   - For events \(A, B\),
     \[
     \Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}.
     \]
   - Summation rule. Let \(E_1, \ldots, E_k\) be events that are disjoint and whose union is the entire sample space (that is, one of the events \(E_i\) will always happen). Then
     \[
     \Pr(A) = \Pr(A \cap E_1) + \cdots + \Pr(A \cap E_k) = \Pr(A|E_1)\Pr(E_1) + \cdots + \Pr(A|E_k)\Pr(E_k).
     \]
   - Bayes rule.
     \[
     \Pr(H|E) = \frac{\Pr(E|H)\Pr(H)}{\Pr(E)}.
     \]
Random variables, expectation, and variance

1. For any random variable $X$, its mean, variance, and standard deviation are

$$E(X) = \sum_z z \Pr(X = z)$$

$$\text{var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

$$\text{stddev}(X) = \sqrt{\text{var}(X)}$$

2. $E(aX + b) = aE(X) + b$.

3. $E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n)$.

4. $\text{var}(aX + b) = a^2 \text{var}(X)$.

5. If $X_1, \ldots, X_n$ are independent, then $\text{var}(X_1 + \cdots + X_n) = \text{var}(X_1) + \cdots + \text{var}(X_n)$.

6. If you toss a coin of bias $p$ until you get heads, the expected number of tosses is $1/p$.

7. Suppose you toss a coin with bias $p$ and code the outcome as either $X = 1$ (heads) or $X = 0$ (tails). We say $X$ has a Bernoulli($p$) distribution. $E(X) = p$ and $\text{var}(X) = p(1-p)$.

Sampling, hypothesis testing, and the central limit theorem

1. If you take a coin of bias $p$ and toss it $n$ times, the number of heads has a binomial($n, p$) distribution.

2. If $X$ has a binomial($n, p$) distribution, then:
   - $E(X) = np$ and $\text{var}(X) = np(1-p)$.
   - With 95% probability, $X$ lies in the range $np \pm 2\sqrt{np(1-p)}$.
   - With at least 95% probability, $X$ lies in the range $np \pm \sqrt{n}$.

3. The normal distribution $N(\mu, \sigma^2)$ has mean $\mu$ and variance $\sigma^2$.

4. If the distribution of $X$ is $N(\mu, \sigma^2)$, then the distribution of $aX + b$ is $N(a\mu + b, a^2\sigma^2)$.

5. If $X$ has a $N(\mu, \sigma^2)$ distribution, then:
   - With 66% probability, $X$ lies between $\mu - \sigma$ and $\mu + \sigma$.
   - With 95% probability, $X$ lies between $\mu - 2\sigma$ and $\mu + 2\sigma$.
   - With 99% probability, $X$ lies between $\mu - 3\sigma$ and $\mu + 3\sigma$.

6. Central Limit Theorem: Suppose $X_1, \ldots, X_n$ are independent with mean $m$ and variance $v$. Then for sufficiently large $n$, the distribution of $(X_1 + \cdots + X_n)/n$ is approximately $N(m, v/n)$. Equivalently, the distribution of $X_1 + \cdots + X_n$ is approximately $N(nm, nv)$.