3 A3. Unbiased Estimation for R@k

Our R@k estimator is unbiased because it estimates the all-correct probability p^k as the probability of drawing k correct samples without replacement. To show this, note that c, the number of correct samples that pass the unit tests, is distributed $\mathrm{Binom}(n, p)$, where p is pass@1, and that our R@k evaluates to 1 when $c \geq k$. Then,

$$
\begin{aligned}
\mathbb{E}_c\left[\frac{\binom{c}{k}}{\binom{n}{k}}\right] &= \sum_{i=k}^{n} \frac{\binom{i}{k}}{\binom{n}{k}} \binom{n}{i} p^i (1-p)^{n-i} \\
&= \frac{1}{\binom{n}{k}} \sum_{i=k}^{n} \binom{i}{k} \binom{n}{i} p^i (1-p)^{n-i} \\
&= \sum_{i=k}^{n} \binom{n-k}{i-k} p^i (1-p)^{n-i} \\
&= p^k \sum_{i=k}^{n} \binom{n-k}{i-k} p^{i-k} (1-p)^{n-k-(i-k)} \\
&= p^k \sum_{j=0}^{n-k} \binom{n-k}{j} p^j (1-p)^{n-k-j} \\
&= p^k.
\end{aligned} \tag{1}
$$