CSE 107 4-23-26

Ex

suppose we have a coin with

$$P(\text{head}) = p$$

which is unknown. How to estimate p?

let $X_1, X_2, \ldots, X_n$ be Bernoulli(p), independent. Define

$$S_n = \frac{X_1 + X_2 + \cdots + X}{n}$$

We call S_n the <u>sample mean</u>, or <u>relative frequency</u>.

Since $E[X_i] = p$, we have

$$E[S_n] = E\left[\frac{X_1 + \cdots + X_n}{n}\right]$$

$$= E\left[\frac{1}{n}X_1 + \cdots + \frac{1}{n}X_n\right]$$

$$= \frac{1}{n}E[X_1] + \cdots + \frac{1}{n}E[X_n]$$

$$= \underbrace{\frac{1}{n}\cdot p + \cdots + \frac{1}{n}\cdot p}_{n}$$

$$= \cancel{n}\cdot\frac{1}{\cancel{n}}\cdot p = p.$$

Thus S_n can serve as estimate for $E[S_n] = p$.

How good is the estimate?

Since X_i are independent, so are $\frac{1}{n} \cdot X_i$ are also indep.

Thus

$$\begin{aligned}
Var(S_n) &= Var\left(\tfrac{1}{n} X_1 + \cdots + \tfrac{1}{n} X_n\right) \\
&= Var\left(\tfrac{1}{n} X_1\right) + \cdots + Var\left(\tfrac{1}{n} X_n\right) \\
&= \tfrac{1}{n^2} \cdot Var(X_1) + \cdots + \tfrac{1}{n^2} Var(X_n) \\
&= \tfrac{1}{n^2} \cdot p(1-p) + \cdots + \tfrac{1}{n^2} \cdot p(1-p)
\end{aligned}$$

$$= n \cdot \frac{1}{n^2} p(1-p)$$

$$= \frac{p(1-p)}{n} \longrightarrow 0 \text{ as } n \to \infty$$

observe $Var(S_n) \to 0$ as $n \to \infty$.

$\therefore$ for large n, S_n is a good approx. of p. ✓

<u>Exercise</u>

Prove that if Y is a discrete r.v. and $Var(Y)=0$ then $Y = \text{const.} = E[Y]$

<u>Note</u>

Even if X_i are <u>not</u> Bernoulli, but still independent, and all with same PMF $p_X(x)$, then the same calculation gives

$$Var(S_n) = \frac{Var(X)}{n} \to 0$$

since $E[S_n] = E[X]$,

we have S_n is a good

approx. to $E[X]$.