Lec8
Thursday, January 27, 2022 8:13 PM

Binomial Distribution $Bin(n, p)$

$X \sim Bin(n, p)$

1) Story: X is # Successes in $\underline{n}$ independent Bern (p) trials.

Prob. Success

2) Sum of indicators R.V's: $X = X_1 + X_2 + \cdots + X_n$, $X_j = \begin{cases} 1 & \text{, if } j\text{th trial is a Success} \\ 0 & \text{, otherwise.} \end{cases}$

where $X_1, \ldots, X_n$ $\underline{\text{i.i.d}} \sim \text{Bern}(p)$

Independent, Identically Distributed

3) $\underline{\text{PMF}}$: $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ $\quad k \in \{1, \ldots, n\}$

Probability Mass Function: Probability of one distinct event happening.

[$X = k$ is the event.]

CDF: Cumulative Distribution Function.

CDF of X

[$X \leq x$ is an event.] $\Rightarrow F(x) = P(X \leq x)$

PMF: (only defined for discrete R.V's).

Discrete: Possible Values $a_1, a_2, \ldots, a_N$ or $a_1, a_2, \ldots$

PMF: $P(X=a_j)=P_j$ for all j. where $P_j \geqslant 0$, $\sum_j P_j = 1$.

Checking the 2nd Condition on $\text{Bin}(n,p)$:

$$\sum_{k=0}^{n} \binom{n}{k} p^{k} (1-p)^{n-k} \underset{\text{Via the Binomial theorem.}}{=} (p+(1-p))^{n} = 1^{n} = 1$$

Now, we want to Prove the following theorem:

$$X \sim \text{Bin}(n,p) \quad , \quad Y \sim \text{Bin}(m,p) \quad \text{independent} \implies X+Y \sim \text{Bin}(n+m,p)$$

1) Immediate from Story.

2) Sum of Indicator R.v's: $X = X_1 + \cdots + X_n$, $Y = Y_1 + \cdots + Y_m \implies X+Y = \sum_{i=1}^{n} X_i + \sum_{j=1}^{m} Y_j$

$\implies$ Sum of $(n+m)$ iid $\sim \text{Bern}(p)$

3) PMF: $P(X+Y=k) \overset{\text{LoTP}}{=} \sum_{j=0}^{k} P(X+Y=k \mid X=j)\, P(X=j)$ $\quad \left[Y+j=k \right]$

$$= \sum_{j=0}^{k} P(\underbrace{Y=k-j}_{\text{Independent}} \mid \underbrace{\cancel{X=j}}) \binom{n}{j} p^{j} (1-p)^{n-j}$$

$$= \sum^{k} \binom{m}{k-j} p^{k-j} (1-p)^{m-k+j} \binom{n}{j} p^{j} (1-p)^{n-j} = p^{k} (1-p)^{m+n-k} \overbrace{\sum^{k} \binom{m}{\cdot} \binom{n}{\cdot}}^{\text{Vandermonde}}$$

$$= \sum_{j=0}^{k} \binom{m}{k-j} p^{k-j} (1-p)^{m-k+j} \binom{n}{j} p^{j} (1-p)^{n-j} = p^{k} (1-p)^{m+n-k} \underbrace{\sum_{j=0}^{k} \binom{m}{k-j}\binom{n}{j}}_{\binom{m+n}{k}}$$

$$= p^{k}(1-p)^{m+n-k}\binom{m+n}{k}$$ = PMF of $Bin(m+n, p)$ k Successes

Hypergeometric distribution:

Have b black, w white distinguishable marbles. Pick random Sample of size n.

Find the distribution of (#white marbles in Sample $= X$)

$X \sim HGeom(w, b, n)$

$$P(X=k) = \frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}} \qquad 0 \leq k \leq w, \quad 0 \leq n-k \leq b$$

→ Hypergeometric distribution.

Comes from Sampling without replacement.

If we have replacement, it'd be reduced to Binomial.

(with Probability ~ # all white marbles).

Checking the normalization Constraint:

$$\overbrace{\sum^{w} \binom{w}{\cdot}\binom{b}{\cdot}}^{\text{Vandermonde}} \qquad \binom{w+b}{}$$

$$\frac{1}{1}\;\overbrace{\sum_{k=0}^{w}\frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}}}^{\text{Vandermonde}} = \frac{\binom{w+b}{n}}{\binom{w+b}{n}} = 1$$