2^{nd} variation on classical r.v

- In a Bernoulli trial with replacement, if $X = \#$ success then $X \sim Bin(n, p)$

- Question : what happens if we have no replacement?

  $\rightarrow$ HypG

Hypergeometric random variable (1)

with replacement
$X \sim \text{Bin}(n, \frac{m}{N})$

Use: Consider the experiment

- Urn containing N balls
- m white balls, $N - m$ black balls
- Sample of size n is drawn without replacement
- Set $X = \#$ white balls drawn

Then

$$X \sim \text{HypG}(n, N, m)$$

Hypergeometric random variable (2)

Notation:

$$X \sim \text{HypG}(n, N, m), \quad \text{for } N \in \mathbb{N}^*, \; m, n \le N, \; \cancel{p \in (0, 1)}$$

State space:

$$\{0, \ldots, n\}$$

Pmf:

$$\mathbf{P}(X = k) = \frac{\overbrace{\binom{m}{k}}^{\text{\# ways to choose } k \text{ W among } m \text{ W}} \overbrace{\binom{N-m}{n-k}}^{\text{\# ways to pick } n-k \text{ B among } N-m}}{\underbrace{\binom{N}{n}}_{\text{\# ways to pick } n \text{ balls among } N \text{ balls}}}, \quad 0 \le k \le n$$

Expected value and variance: Set $p = \frac{m}{N}$. Then

$$\mathbf{E}[X] = np, \qquad \mathbf{Var}(X) = np(1-p)\left(\frac{N-n}{N-1}\right)$$

Hypergeometric and binomial

Proposition 14.

Let

- $X \sim \mathrm{HypG}(n, N, m)$,
- Recall that $p = \frac{m}{N}$

$\mathrm{HypG}(n, N, m) \approx \mathrm{Bin}(n, \frac{m}{N})$
if N, m are large

Hypothesis:

$$n \ll m, N, \qquad i \ll m, N$$

Then

$$\mathbf{P}(X = i) \simeq \binom{n}{i} p^i (1-p)^{n-i}$$

Proof

Expression for $\mathbf{P}(X = i)$:

$$\begin{aligned}
\mathbf{P}(X = i) &= \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}} \\
&= \frac{m!}{(m-i)!i!} \frac{(N-m)!}{(N-m-n+i)!(n-i)!} \frac{(N-n)!n!}{N!} \\
&= \binom{n}{i} \prod_{j=0}^{i-1} \frac{m-j}{N-j} \prod_{k=0}^{n-i-1} \frac{N-m-k}{N-i-k}
\end{aligned}$$

Approximation: If $i, j, k \ll m, N$ above, we get

$$\mathbf{P}(X = i) \simeq \binom{n}{i} p^i (1-p)^{n-i}$$

Example: electric components (1)

Situation: We have

- Lots of electric components of size 10
- We inspect 3 components per lot
  $\hookrightarrow$ Acceptance if all 3 components are non defective
- 30% of lots have 4 defective components
- 70% of lots have 1 defective component

Question:
What is the proportion of ~~rejected~~ lots?
accepted

Modeling

Define

A = "lot accepted"

70% of lots

L_1 = " we pick a lot with 1 defective "

L_4 = " " " " " 4 defective "

30% of lots

We wish to compute

$$\mathbb{P}(A) \overset{\text{Bayes I}}{=} \mathbb{P}(A \mid L_1)\,\mathbb{P}(L_1) + \mathbb{P}(A \mid L_4)\,\mathbb{P}(L_4)$$

$$\mathbb{P}(A) = 0.7 \times \mathbb{P}(A \mid L_1) + 0.3\ \mathbb{P}(A \mid L_4)$$

$$\mathbb{P}(A) = 0.7 \times \mathbb{P}(A|L_1) + 0.3\,\mathbb{P}(A|L_4)$$

Claim

$HypG(n, N, m)$

trials (n), total # of balls (N), # w balls (m)

$$\mathbb{P}(A|L_1) = \mathbb{P}(X_1 = 0)$$

where X_1 = # defective picked

$$\Rightarrow X_1 \sim HypG(3, 10, 1)$$

$X_4 \sim Hyp(3, 10, 4)$

$$\mathbb{P}(A|L_1^{\,4}) = \mathbb{P}(X_{1_4} = \overset{=0}{k}) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}$$

$$= \frac{\binom{1}{0}\binom{9}{3}}{\binom{10}{3}} \rightarrow \frac{\binom{4}{0}\binom{6}{3}}{\binom{10}{3}}$$

Conclusion

$$\mathbb{P}(A) = .7 \times \frac{\binom{1}{0}\binom{9}{3}}{\binom{10}{3}} + .3 \times \frac{\binom{4}{0}\binom{6}{3}}{\binom{10}{3}}$$

$$\boxed{\mathbb{P}(A) \simeq 54\%}$$

Example: electric components (2)

Events: We define

- A = Acceptance of a lot
- L_1 = Lot with 1 defective component drawn
- L_4 = Lot with 4 defective components drawn

Conditioning: We have

$$\mathbf{P}(A) = \mathbf{P}(A|\, L_1)\, \mathbf{P}(L_1) + \mathbf{P}(A|\, L_4)\, \mathbf{P}(L_4)$$

and

$$\mathbf{P}(L_1) = .7, \qquad \mathbf{P}(L_4) = .3,$$

Example: electric components (3)

Hypergeometric random variable: We check that

$$\mathbf{P}(A|\, L_1) = \mathbf{P}(X_1 = 0), \quad \text{where} \quad X_1 \sim \mathrm{HypG}(3, 10, 1)$$

Thus

$$\mathbf{P}(A|\, L_1) = \frac{\binom{1}{0}\binom{9}{3}}{\binom{10}{3}}$$

Conclusion:

$$\mathbf{P}(A) = \frac{\binom{1}{0}\binom{9}{3}}{\binom{10}{3}} \times 0.7 + \frac{\binom{4}{0}\binom{6}{3}}{\binom{10}{3}} \times 0.3 = 54\%$$

Outline

Expectation of sums

Proposition 15.

Let

- $\mathbf{P}$ a probability on a sample space S
- $X_1, \ldots, X_n : S \to \mathbb{R}$ n random variables

Hypothesis: S is countable, i.e

$$S = \{s_i; i \geq 1\}$$

Then

$$\mathbf{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbf{E}[X_i]$$

Rmk If $X \sim Bin(n,p)$ we can decompose X as

$$X = \sum_{i=1}^{n} X_i, \quad \text{where}$$

$$X_i = \begin{cases} 1 & \text{if success at } i\text{-th trial} \\ 0 & \text{otherwise} \end{cases}$$

Morever

$X_i \sim B(p)$

easy way to compute $E[X]$ if $X \sim Bin$

$$\Rightarrow \boxed{E[X]} = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \overbrace{E[X_i]}^{=p}$$

$$\boxed{= np}$$