$\mathbb{P}(A|B) = \mathbb{P}(A)$

Common univariate distributions

Discrete distributions: Negative Binominal

$X \sim$Neg-Binom(r, p), $0 < p < 1$

- pmf given by

$r = 1$

$$f_X(x) = f_X(x|r,p) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \quad x = r, r+1, \ldots,$$

- Motivation: distribution for the number of independent Bernoulli(p) trials needed to obtain r successes

- $Y = X - r$ (number of failures prior to the rth success) also common

$$f_Y(y|r,p) = \binom{y+r-1}{y} p^r (1-p)^y, \quad y = 0, 1, \ldots,$$

- Showing that these probabilities sum to 1 is not easy (next slide)

- Be careful: both r.v.s X and Y (different) are called "negative binomial"

- Mean: $\mathrm{E}Y = \dfrac{r(1-p)}{p}$ and hence $\mathrm{E}X = \mathrm{E}Y + r = \dfrac{r}{p}$

- Variance: $\mathrm{Var}(Y) = \dfrac{r(1-p)}{p^2} = \mathrm{Var}(X)$

- $M_Y(t) = \mathrm{E}e^{tY} = \left[\dfrac{p}{1-(1-p)e^t}\right]^r$, $t < -\log(1-p)$,

  $M_X(t) = \mathrm{E}e^{t(Y+r)} = \mathrm{E}e^{rt}e^{tY} = e^{rt}M_Y(t)$

Common univariate distributions

Discrete distributions: Negative Binomial (cont'd)

To show probabilities sum to 1:

1. Newton's negative binomial formula : if $\alpha < 0$ and $|x| < 1$,

$$(1+x)^\alpha \;=\; \sum_{k=0}^{\infty} g^{(k)}(0)\frac{(x-0)^k}{k!} \;=\; \sum_{k=0}^{\infty} \binom{\alpha}{k}^{*} x^k, \quad \binom{\alpha}{k}^{*} \equiv \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}$$

Taylor expanding $g(x) = (1+x)^\alpha$ around 0: $g^{(0)}(0) = g(0) = 1$, $g^{(1)}(0) = g'(0) = \alpha$

2. for integers $r \geq 1$ and $k \geq 0$, note that

$$\binom{-r}{k}^{*}(-1)^k = (-1)^k \frac{(-r)(-r-1)\cdots(-r-k+1)}{k!} = \frac{(r)(r+1)\cdots(r+k-1)}{k!} = \binom{r+k-1}{k}$$

$$\begin{aligned}
\sum_{y=0}^{\infty} f_Y(y) = \sum_{y=0}^{\infty} \binom{y+r-1}{y} p^r (1-p)^y \;&=\; \sum_{y=0}^{\infty} \binom{-r}{y}^{*} (-1)^y p^r (1-p)^y \\
&=\; p^r \sum_{y=0}^{\infty} \binom{-r}{y}^{*} (p-1)^y \\
&=\; p^r [1 + (p-1)]^{-r} = 1
\end{aligned}$$

Show $M_Y(t) = Ee^{tY} = \left[\dfrac{p}{1-(1-p)e^t}\right]^r$ for $t < -\log(1-p)$

Common univariate distributions

Discrete distributions: Geometric

$X \sim$Geom(p), $0 < p < 1$

- special case of Negative Binomial($r = 1, p$)

- Motivation: distribution for the number of independent Bernoulli(p) trials needed to obtain 1st success

- pmf given by

$$\mathbb{P}(X=x) = f_X(x) = f_X(x|p) = p(1-p)^{x-1}, \quad x = 1, 2, 3, \ldots,$$

- Mean: $\mathrm{E}X = \dfrac{1}{p}$

- Variance: $\mathrm{Var}(X) = \dfrac{1-p}{p^2}$

- $M_X(t) = Ee^{tX} = \dfrac{pe^t}{1-(1-p)e^t}$ for $t < -\log(1-p)$

Recall: $X \sim$ Neg-Bin(r,p)

$$M_X(t) = \frac{e^{rt} p^r}{(1-(1-p)e^t)^r}$$

$$\xRightarrow{r=1} M_X(t) = \frac{e^t p}{1-(1-p)e^t}$$

Common univariate distributions

Discrete distributions: Geometric (cont'd)

1. We've seen the cdf of a Geom(p) random variable X; it's relatively simple.

$$X \sim \text{Geo}(p) \Rightarrow F_X(x) \overset{\text{def}}{=} \mathbb{P}(X \leq x)$$

$$1 - F_X(x) = \mathbb{P}(X > x) = (1-p)^x \quad (*)$$

$$F_X(x) = 1 - (1-p)^x \text{ for } x = 1, 2, \cdots$$

2. The Geometric distribution has the famous "memoryless" property: for any integer $x_0 \geq 0$.

$$P(X = x_0 + x | X > x_0) = P(X = x)$$

Interpretation: The conditional distribution of the remaining waiting number of trials until a 1st success, given that I've already waited x_0 trials, is the same as the original distribution of the number of trials until 1st success.

Given that each trial is an independent Bernoulli trial, this does make sense: whether I start counting trials at the beginning or I start counting trials after x_0 trials without success, the distribution of the remaining number of trials needed until a 1st success should be the same.

$$\mathbb{P}(X = x + x_0 | X > x_0) = \frac{\mathbb{P}(X = x + x_0 \cap X > x_0)}{\mathbb{P}(X > x_0)} = \frac{\mathbb{P}(X = x + x_0)}{\underbrace{\mathbb{P}(X > x_0)}_{(*)}}$$

$$= \frac{p\cancel{(1-p)}^{x + \cancel{x_0} - 1}}{\cancel{(1-p)^{x_0}}} = \underbrace{p(1-p)^{x-1}}_{= \mathbb{P}(X = x)} \quad \blacksquare$$

3. Example: Testing newly manufactured widgets with $p = 0.01$ probability that a given widget fails a functionality test, what's the probability of running at least 50 units without a test failure?

$\longrightarrow X =$ # of trails (tests) until the 1st failure

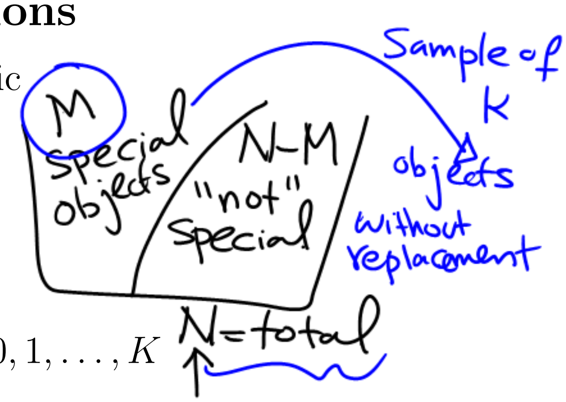
$$X \sim \text{Geom}(.01)$$

$$\mathbb{P}(X > 50) = (1-p)^x = (1 - .01)^{50} \approx .605$$

$$\mathbb{E}(X) = \frac{1}{p} = \underline{\underline{100}}$$

Common univariate distributions

Discrete distributions: Hypergeometric



X ∼Hypergeometric(N, M, K) (integers N, M, K)

- pmf given by

$$\mathbb{P}(X=x) = f_X(x) = \binom{M}{x}\binom{N-M}{K-x}\Big/\binom{N}{K}, \quad x = 0, 1, \ldots, K$$

- Motivation: Choose K objects without replacement from a total population of size N which contains M "special" objects. X is the number of "special" objects among the K chosen.

$$\mathbb{P}(X=x) = \frac{M!}{x!\,(M-x)!}\,\frac{(N-M)!}{(K-x)!\,(N-M-K+x)!}\,\frac{K!\,(N-K)!}{N!}$$

pmf of $X \sim \text{Hyper}(N, M, K)$ total

for $x = 0, 1, 2, \ldots, K$

- Must have $0 \le x \le K$, $x \le M$, and $K - x \le N - M$ in $f_X(x)$

  Typically, $N > 2M$ and $M > K$ so only the condition $0 \le x \le K$ matters

- Mean: $\mathrm{E}X = KM/N$

  size of your sample × % of special object

$$\mathbb{P}(X=x) = \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}} \quad \text{for } x = 0, 1, 2, \ldots, K$$

special

- Variance: $\mathrm{Var}(X) = \dfrac{KM(N-M)(N-K)}{N^2(N-1)}$

To derive mean:

$$\mathrm{E}X = \sum_{x=0}^{K} x \overbrace{\binom{M}{x}\binom{N-M}{K-x}\Big/\binom{N}{K}}^{\mathbb{P}(X=x)} = \sum_{x=1}^{K} x \frac{M!}{x!(M-x)!}\binom{N-M}{K-x}\frac{K!(N-K)!}{N!}$$

def $\sum_{x=0}^{K} x\,\mathbb{P}(X=x)$

$x! \to x(x-1)!$ $\quad K! \to K(K-1)!$

$$= \sum_{x=1}^{K} \frac{\cancel{x}\,M\,(M-1)!}{\cancel{x}(x-1)!}\binom{N-M}{K-x}\frac{K(K-1)!\,(N-K)!}{N(N-1)!}$$

$$= \frac{KM}{N}\sum_{x=1}^{K} \frac{(M-1)!}{(M-x)!\,(x-1)!}\binom{N-M}{K-x}\frac{(K-1)!\,(N-K)!}{(N-1)!}$$

$Y \sim \text{HyperGeo}(N-1, M-1, K-1)$ $\quad y = x-1 \quad \frac{KM}{N} \quad \sum_{y=0}^{K-1} \frac{(M-1)!}{(M-1-y)!\,(N-1)!} \binom{N-1-(M-1)}{K-1-(x-1)} \frac{(K-1)!\,(N-1-(K-1))!}{(N-1)!} = 1$

Common univariate distributions

Discrete distributions: Hypergeometric and Binomial

1. Hypergeometric: sampling *without* replacement

   i.e., choose x special objects in a size K sample from a collection where M objects are "special" & $N - M$ are not

2. Binomial: sampling *with* replacement

   i.e., choose x special objects in a size n sample, where each selected item of the sample has probability p of being a special object

   $\text{Bin}\left(n = K,\ p = \frac{M}{N}\right)$

Suppose X ∼Hypergeometric(N, M, K) and let $M/N = p$ be the proportion of "special" objects and let $K = n$ be the sample size. Then,

- $EX = \overset{n}{K}\,\overset{p}{M/N} = np$
- $\text{Var}(X) = \dfrac{KM(N-M)(N-K)}{N^2(N-1)} = np(1-p)\dfrac{N-n}{N-1}$
- the factor $(N-n)/(N-1)$ in Var(X) is the finite sample correction factor

Note: $N \geq 20K$

sample

$N \geq 20n$

Now fix $K = n$ and $p = M/N$ as above, and let $N \to \infty$. Then,

$$\lim_{N\to\infty} P(X = x) = \lim_{N\to\infty} \binom{M}{x}\binom{N-M}{K-x} \Big/ \binom{N}{K} = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, \ldots, n$$

Result: hypergeometric tends to binomial distribution in large population size N