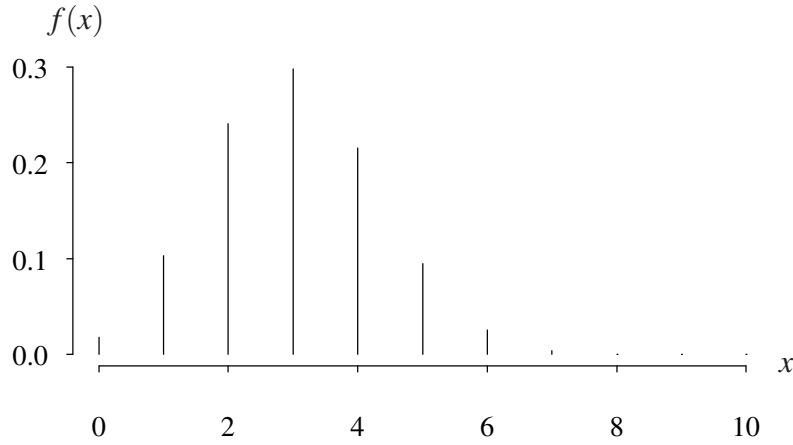


Hypergeometric distribution (from <http://www.math.wm.edu/~leemis/chart/UDR/UDR.html>)

The shorthand $X \sim \text{hypergeometric}(n_1, n_2, n_3)$ is used to indicate that the random variable X has the hypergeometric distribution for some nonnegative integer parameters n_1, n_2 , and n_3 , where $n_1, n_2 \in \{0, 1, 2, \dots, n_3\}$. A hypergeometric random variable X for parameters n_1, n_2 , and n_3 has probability mass function

$$f(x) = \frac{\binom{n_1}{x} \binom{n_3 - n_1}{n_2 - x}}{\binom{n_3}{n_2}}.$$

for $x = \max\{0, n_1 + n_2 - n_3\}, \dots, \min\{n_1, n_2\}$. The hypergeometric distribution is used for sampling without replacement from a finite population of items. More specifically, a hypergeometric random variable X is the number of defective items in a sample of size n_2 items drawn at random and without replacement from a lot of n_3 items which contains n_1 defective items. Applications include acceptance sampling from quality control and animal population size estimation using tagging with capture/recapture. The probability mass function for $n_1 = 15$, $n_2 = 10$, and $n_3 = 50$ is illustrated below.



The cumulative distribution function on the support of X is

$$F(x) = P(X \leq x) = \sum_{k=0}^x \frac{\binom{n_1}{k} \binom{n_3 - n_1}{n_2 - k}}{\binom{n_3}{n_2}} \quad x = \max\{0, n_1 + n_2 - n_3\}, \dots, \min\{n_1, n_2\}.$$

The moment generating function (from Wikipedia) of X is

$$M_X(t) = \frac{\binom{n_3 - n_1}{n_2}}{\binom{n_3}{n_2}} {}_2F_1(-n_2, -n_1, n_3 - n_1 - n_2 + 1, e^t),$$

where ${}_2F_1$ is the hypergeometric function defined by

$${}_2F_1(a, b, c, z) = 1 + \frac{ab}{c} \cdot \frac{z}{1!} + \frac{a(a+1)b(b+1)}{c(c+1)} \cdot \frac{z^2}{2!} + \frac{a(a+1)(a+2)b(b+1)(b+2)}{c(c+1)(c+2)} \cdot \frac{z^3}{3!} + \dots$$

The population mean, variance, and skewness of X are

$$E[X] = \frac{n_1 n_2}{n_3} \quad V[X] = \frac{n_2 n_1 (n_3 - n_1)(n_3 - n_2)}{n_3^2 (n_3 - 1)}$$

$$E \left[\left(\frac{X - \mu}{\sigma} \right)^3 \right] = \frac{(n_3 - 2n_1)(n_3 - 1)^{1/2}(n_3 - 2n_2)}{[n_2 n_1 (n_3 - n_1)(n_3 - n_2)]^{1/2} (n_3 - 2)}.$$

APPL verification: The APPL statements

```
X := HypergeometricRV(n3, n1, n2);  
Mean(X);  
Variance(X);  
Skewness(X);
```

return complicated expressions for the population mean, variance, and skewness.