STAT 2857A – Lecture 11 Examples and Exercises

Solutions

Example 11.1

a) The random variable Y follows a hypergeometric distribution because:

- The population consists of a finite number of individuals (accounts), N.
- M individuals can be classified as successes (accounts with errors) and $N-M$ as failures (accounts without errors).
- A sample of n individuals is chosen *without replacement* so that each sample is equally likely.

In this case, $N = 200$, $M = 20$, and $n = 10$. So

$$Y \sim \text{Hypergeometric(10,20,200)}.$$

b) The probability that the accountant selects 2 accounts with errors is

$$\begin{aligned} P(Y=2) &= h(2; 10, 20, 200) \\ &= \frac{\binom{20}{2}\binom{200-20}{10-2}}{\binom{200}{10}} \\ &= 0.176025. \end{aligned}$$

The probability the accountant chooses 2 accounts with errors is 0.1760.

c) The mean and variance of Y are:

$$\begin{aligned} E(Y) &= \frac{nM/N}{=}\frac{10(20)}{200} = 1.00 \\ V(Y) &= \left(\frac{N-n}{N-1}\right)\left(\frac{nM}{N}\right)\left(1-\frac{M}{N}\right) \\ &= \left(\frac{200-10}{200-1}\right)\left(\frac{10(20)}{200}\right)\left(1-\frac{20}{200}\right) \\ &= .8593 \end{aligned}$$

Example 11.2

a) According to your textbook, the binomial approximation is appropriate because $n/N = 10/200 = .05$.

b) The approximate distribution is

$$Y \dot{\sim} \text{Binomial}(10, .10).$$

c) The approximate probability is

$$\begin{aligned} P(Y = 5) &\approx \binom{10}{5} .10^2 (1 - .10)^8 \\ &= 0.19371. \end{aligned}$$

Note that this is close, but not exactly equal to, the exact probability computed in part a). Whether or not this approximation is good enough will depend on the situation.

d) The approximate mean is

$$\mu_Y \approx 10(.10) = 1.$$

The approximate variance is

$$\sigma_Y^2 \approx 10(.15)(1 - .10) = .9000.$$

In this case, the mean of the approximate distribution is equal to the exact mean. The variance is close, but not exact. Again, whether or not this approximation is good enough will depend on the situation.

Example 11.3

a) Note that the probability that any one account sampled has errors is $20/200 = .10$. Hence

$$Z \sim \text{Neg. Binomial}(2, .1).$$

b) The probability that the accountant samples 10 accounts without errors before reaching 2 with errors is

$$\begin{aligned} P(Z = 5) &= nb(2; 10, .10) \\ &= \binom{10 + 2 - 1}{2 - 1} .1^2 (1 - .1)^{10} \\ &= 0.03835 \end{aligned}$$

The probability that the accountant samples 10 accounts without errors before reaching 2 with errors is 0.0384.

c) The mean and variance are

$$
\begin{aligned}
E(Z) &= \frac{2(1-.1)}{.1} = 18 \\
V(Z) &= \frac{2(1-.1)}{.1^2} = 180
\end{aligned}
$$

Exercise 11.1

a) Let X by the number of red smarties in a package. Then X has a binomial distribution because the number of trials (Smarties per package) is fixed and the probability of success (a candy being red) is also fixed. The distribution is

$$X \sim \text{Binomial}(30, .25).$$

The mean and variance of X are

$$
\begin{aligned}
\mu_X &= 30(.25) = 7.5 \\
\sigma_X^2 &= 30(.25)(1-.25) = 5.625.
\end{aligned}
$$

Using the calculator available on stattrek

$$P(X > 10) = 0.10573.$$

b) Let Y be the number of of red smarties you draw. Sampling occurs without replacement, so the distribution is hypergeometric. Specifically,

$$Y \sim \text{Hypergeometric}(5, 8, 30).$$

The mean and variance of Y are

$$
\begin{aligned}
E(Y) &= \frac{nM/N}{=}\frac{5(8)}{30} = 1.33333 \\
V(Y) &= \left(\frac{N-n}{N-1}\right)\left(\frac{nM}{N}\right)\left(1-\frac{M}{N}\right) \\
&= \left(\frac{30-5}{30-1}\right)\left(\frac{5(8)}{30}\right)\left(1-\frac{8}{30}\right) \\
&= 0.84291
\end{aligned}
$$

The probability that this number is less than three is

$$P(Y < 3) = P(Y = 0) + P(Y = 1) + P(Y = 2)$$

Using the calculator available at stattrek

$$P(Y < 3) = 0.89802.$$

c) Let Z denote the number of boxes you must open. In this case, the probability of success is fixed but the number of trials (boxes you must open) is random. Hence, the distribution is negative binomial. The probability that a box contains no red smarties is

$$P(X = 0) = .75^{30} = 0.000178582.$$

Hence,

$$Z \sim \text{Neg. Binomial}(1, 0.000178582).$$

This is the same as saying that

$$Z \sim \text{Geometric}(0.000178582).$$

The mean and variance of Z are

$$\mu_Z = \frac{1(1 - 0.000178582)}{0.000178582} = 5598.666$$
$$\sigma_Z^2 = \frac{1(1 - 0.000178582)}{0.000178582^2} = 31,350,656.$$

Again, we can use the calculator available at stattrek. However, we have to keep in mind that the calculator uses a slightly different version of the negative binomial. Instead of considering the number of failures, z, that occur before r successes it considers the number of trials, $x = z + r$, to reach the z successes. In this case, we are interested in the probability of opening 5000 boxes which means that the first $z = 4999$ all contain red smarties and the last one contains no red smarties. The probability is

$$P(Z = 4999) = .00007.$$