Rohit Wason
Spring 2021

Math 560 Homework (#5, Binomial)

Problem 1.

Solution. □

Problem 2. Population data on StatVillage (a hypothetical 128-block village in Canada) is given in the tab-delimited data file `StatVillage.txt`. The variables are listed in the first line of the data file, and information about the variables included in the file is given in the file `codesForStatVillage.txt`. Use `R` or other computer software to answer the following questions:

```
population = read.table(file ="StatVillage.txt",
  header = TRUE)
```

Solution. (a) The variable labeled `TOTINCH` gives the total household income. Determine the proportion of households in this population with a total household income greater than 100,000.

```
nrow(population[population$TOTINCH>100000,])/
  nrow(population)
```

$$p = \frac{114}{1024} \approx 0.1113$$

(b) If 100 households are selected at random with replacement from this population, what is the probability that at least 10 of the households in the sample will have a total household income greater than 100,000? Compute the exact answer, rounded to at least 4 decimal places.

Adding all probabilities from $X = 10, 11, \ldots, 100$ in R

```
prob=0;
for (n in c(10:100))
  prob = prob + dbinom(n, size=100, prob=0.1113);
prob;
```

we get $\boxed{P(X \geq 10) = 0.6868}$

(c) If 100 households are selected at random with replacement from this population, then let X be the number of households in the sample with income above $100,000$. What is the mean and the standard deviation of the sampling distribution of X?

X follows $\approx B(n,p) = B(100, 0.11)$ from above. According to CLT, this also $\approx N(np, \sqrt{np(1-p)}) = N(11.13, \sqrt{11.13(0.8887)}) = N(11.13, 3.1450)$

Or $\mu = 11.13, \sigma = 3.1450$

(d) Use the central limit theorem with a continuity correction to approximate the probability computed in part (b)

Using $\mu = 11.13, \sigma = 3.1450$ from (c), for $Z = \frac{X-11.13}{3.1450}$, the required probability,

$$\begin{aligned} P(X \geq 10) & \\ &= P(\frac{X - 11.13}{3.1450} \geq \frac{10 - 11.13}{3.1450}) \\ &= P(Z \geq \frac{-1.13}{3.1450} = -0.3593) \\ &= 1 - P(Z < -0.3593) \\ &= 1 - 0.3632 \\ &= \boxed{0.6368} \end{aligned}$$

□