MATH10282 Introduction to Statistics
Semester 2, 2020/2021
Example Sheet 6

1. Based on long experience, an airline has found that on average 6% of people who have bought tickets for a flight on a particular route do not show up for the flight. (Some may be late while others may not be able to travel.) Suppose that the airline overbooks this flight by selling 267 tickets for an aircraft with only 255 seats. Let the random variable Y denote the number of people that show up for the flight having bought a ticket.

(i) What is the probability distribution of Y? State any assumptions that you have made. What are $\mathrm{E}(Y)$ and $\mathrm{Var}(Y)$?

(ii) Is it appropriate to approximate the distribution of Y with a normal distribution?

(iii) What is the *approximate* probability that the flight will take off with between 248 and 255 passengers?

(iv) What is the *approximate* probability that everybody who shows up for the flight will have a seat available for them?

(v) What is the *approximate* probability that between six and seven percent of passengers with tickets do not show up for the flight?

2. (i) Show that if $n \geq 9 \max\left\{\frac{1-p}{p}, \frac{p}{1-p}\right\}$ then

$$\frac{n - np}{\sqrt{np(1-p)}} \geq 3 \quad \text{and} \quad \frac{-np}{\sqrt{np(1-p)}} \leq -3\,.$$

(ii) Hence show that if $n \geq 9 \max\left\{\frac{1-p}{p}, \frac{p}{1-p}\right\}$ and $Y \sim N[np, np(1-p)]$, then

$$\mathrm{P}(0 \leq Y \leq n) \geq 0.997\,,$$

i.e. prove Proposition 3.3 from the lectures.

(iii) Using the rule of thumb from lectures, find the minimum n for which the normal approximation to the Binomial is valid in the case $p = 0.1$, and find the range of p for which it is valid when $n = 100$.

3. Let $X_1, \ldots, X_n$ be a random sample of size $n = 25$ observations from $N(112, 12^2)$.

(i) What is the sampling distribution of $\bar{X}_{25}$?

(ii) Find the probability that $\bar{X}_{25} > 115$.

(iii) Find the probability that $\bar{X}_{25}$ deviates from μ by no more than 1.

4. Two separate, independent random samples of sizes $n_1 = 10$ and $n_2 = 15$ are taken from a $N(20, (\sqrt{3})^2)$ distribution. What is the probability that the means of the two samples differ (in absolute value) by more than 0.3?

5. It is believed that 65% of voters will support Candidate A in an upcoming election. Data are collected for two independent random samples, each comprising 200 voters. Calculate the probability that the proportions supporting Candidate A in the two samples will differ by less than 0.1.

6. (i) Let $X_1, \ldots, X_n$ be a random sample of size n from a $\text{Po}(\lambda)$ distribution, where λ is unknown. What are the mean and variance of the sampling distribution of the estimator $\hat{\lambda} = \bar{X}_n$? What is the approximate distribution of $\hat{\lambda}$ for large n?

(ii) Now suppose that a random sample of size $n = 100$ is drawn from a Poisson distribution with parameter $\lambda = 36$. What is the *approximate* probability that the value of the sample mean is greater than 35.0 but less than 37.0?

7. Let $X_1, \ldots, X_n$ be a random sample of size $n = 10$ from $N(48, 36)$.

(i) What is the probability that the sample variance lies between 25 and 60?

(ii) How large a sample should we take so that $\text{P}(S^2 > 20) = 0.9$?

[Hint: use the `pchisq` function in R to find chi-square probabilities.]

8. Let $X_1, \ldots, X_n$ be a random sample of size n from a population with known mean $\mu = 0$ and unknown variance σ^2. Find the constant k such that

$$\hat{\sigma}^2 = k \sum_{i=1}^{n} X_i^2$$

is an unbiased estimator of σ^2. If the X_i were independent $N(0, \sigma^2)$ random variables, what would be the sampling distribution of an appropriately scaled version of $\hat{\sigma}^2$?

9. Let $X_1, \ldots, X_n$ be a random sample of size n from a $U(0, \theta)$ distribution, where the parameter θ is unknown. For what value of the constant k is the estimator

$$\hat{\theta} = k\bar{X}_n$$

an unbiased estimator of θ? For this value of k, what is $\text{Var}(\hat{\theta})$? What happens to $\text{Var}(\hat{\theta})$ as $n \to \infty$?

10. Let $X_1, \ldots, X_n$ be a sample of n independent $\text{Bi}(m, p)$ random variables, where p is unknown. Consider the estimator of p given by

$$\hat{p} = \frac{\bar{X}_n + 1}{n + 2}.$$

Find the bias and variance of $\hat{p}$ and say what happens to the values of these quantities as $n \to \infty$.