Statistics 6 Tutorial: Sampling

Basic Mastery Questions

1. A random sample of 250 adult men undergoing a routine medical inspection had their heights (x cm) measured to the nearest centimeter, and the following data was obtained $\sum x = 43\,205$ and $\sum x^2 = 7\,469\,107$.

(a) Calculate an unbiased estimates of the population mean and variance.
(b) Describe the process of how the random sample could be obtained, assuming that the random sample was from a population of 1000 adult men, context being from a particular company's routine medical check-up.

Solution

(a) $$\bar{x} = \frac{\sum x}{n} = \frac{43205}{250} = 172.82 \approx 173$$

$$s^2 = \frac{1}{n-1}\left[\sum x^2 - \frac{\left(\sum x\right)^2}{n}\right] = \frac{1}{249}\left[7469107 - \frac{(43205)^2}{250}\right] = 9.7144 \approx 9.71$$

(b) The **random sample** could be obtained as follows:

(1) Create a complete list of all members of the population from the human resources department.

(2) Assign each member a unique number (1,2,3, … , N)

(3) Choose a random sample by generating 250 random numbers (via a GC, computer or handheld calculator). Note the numbers and collect data of the men assigned those selected numbers.

2. (i) The time taken for a group of people to complete a survey is normally distributed with mean 10 minutes and standard deviation 2 minutes.
5 individuals are selected at random. For this sample, calculate the probability that the mean time for completing the survey is shorter than 9 minutes.

(ii) A random sample of 40 observations is selected from a population with mean 20 and standard deviation 2.9. Given that $\bar{X}$ is the sample mean, find

(a) $P\left(\bar{X} > 18.5\right)$,

(b) $P\left(14 < \bar{X} \leq 19.5\right)$,

(c) the value of a if $P\left(\bar{X} > a\right) = 0.05$.

Solution

(i) Let T be the time taken for a person to complete a survey. Since T is normally distributed, then $\overline{T}$ is normally distributed as well.

$\overline{T} \sim \mathrm{N}\left(10, \frac{2^2}{5}\right)$

$\mathrm{P}\left(\overline{T} < 9\right) = 0.131776... = 0.132$ (3 s.f.).

(ii) *{Note that the population distribution is not known. But n is large so CLT applies.}*

By CLT, $\overline{X} \sim \mathrm{N}\left(20, \frac{2.9^2}{40}\right)$ approximately

(a) $\mathrm{P}(\overline{X} > 18.5) = 0.9994647 = 0.999$ (3 s.f.)

(b) $\mathrm{P}(14 < \overline{X} \leq 19.5) = 0.1377595 = 0.138$ (3 s.f.)

(c) $\mathrm{P}(\overline{X} > a) = 0.05 \Rightarrow \mathrm{P}(\overline{X} < a) = 0.95$

$a = 20.8$ (3 s.f.)

$a = 20.8$
