

STA 6207 – Homework 1 – Fall 2022

Part 1: Sampling from a Normal Distribution

Obtain random samples from a Normal with mean $\mu = 100$, $\sigma = 20$ of sizes $n = 100, 1000, 10000, 100000$.

- On a single page (2 rows, 2 columns) give the histograms on the same set of bins, with a normal density superimposed on each. Comment on the approximation accuracy.
- For each sample size, give the mean, standard deviation, and the following percentiles (2.5, 25, 50, 75, 97.5). Compare these with the theoretical values.

Part 2: Sampling Distributions of Statistics – Normal Data

Obtain 10000 samples of: $n_1 = 25$ from $N(\mu_1 = 60, \sigma_1 = 10)$ and $n_2 = 15$ from $N(\mu_2 = 50, \sigma_2 = 8)$.

For each sample obtain the following quantities:

$$\bar{y}_1, s_1, \bar{y}_2, s_2, X_1^2 = \frac{(n_1 - 1)s_1^2}{\sigma_1^2}, X_2^2 = \frac{(n_2 - 1)s_2^2}{\sigma_2^2}, t_1 = \frac{\bar{y}_1 - \mu_1}{s_1/\sqrt{n_1}}, t_2 = \frac{\bar{y}_2 - \mu_2}{s_2/\sqrt{n_2}}, F = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2}$$

- Give the theoretical sampling distributions of all of the quantities (except s_1 and s_2 , which are not “standard” distributions).
- Plot histograms of X_1^2, t_1, F and superimpose their theoretical densities.
- For the three sample quantities in the previous part, give the mean, standard deviation, and the following percentiles (2.5, 25, 50, 75, 97.5). Compare these with the theoretical values.

Part 3: Sampling Distributions of Statistics – NHL BMIs and Marathon Velocities

Part 3a: NHL Body Mass Indices (BMI)

For the population of 2013/14 National Hockey League players, obtain their mean and standard deviation. Give a histogram and superimposed normal density. Does a Normal distribution seem to be a reasonable model?

Compute the population mean μ and variance σ^2 .

- Obtain 10000 random samples of $n = 25$ from this population, and save the following quantities:
 $\bar{y}, s, \frac{(n-1)s^2}{\sigma^2}$
- Obtain the mean, standard deviation, and the following percentiles (2.5, 25, 50, 75, 97.5) for the sample mean and scaled variance. Compare these with the theoretical values.
- Plot the sampling distribution of the sample mean and the scaled variance. Superimpose their theoretical densities assuming normality.

Part 3b: Rock and Roll Marathon

For the population of 2015 Rock and Roll Marathon Participants, obtain their mean and standard deviation by gender. Give histograms and superimposed normal densities. Does a Normal distribution seem to be a reasonable model? Compute the population mean μ and variance σ^2 by gender.

- Obtain 10000 random samples of $n_F = n_M = 20$ from this population, and save the following quantities:

$$\bar{y}_F, s_F, \bar{y}_M, s_M, \frac{(n_F - 1)s_F^2}{\sigma_F^2}, \frac{\left(\frac{s_F^2}{\sigma_F^2}\right)}{\left(\frac{s_M^2}{\sigma_M^2}\right)}$$

- Obtain the mean, standard deviation, and the following percentiles (2.5, 25, 50, 75, 97.5) of the mean and scaled variance for Females, as well as the F ratio. Compare these with the theoretical values, assuming normality.

Part 4: Size and Power of Likelihood Ratio, Wald, and Score Tests

Generate 10000 Pseudo-Random Samples from Poisson Distributions. Test $H_0: \mu = 3$ (2-sided) based on LR, Wald, and Score Tests. Give the (Empirical) Power of each Test for the following settings. A) $n = 20, \mu = 3$. B) $n = 20, \mu = 2.57$. C) $n = 100, \mu = 3$. D) $n = 100, \mu = 2.57$.