The aim of this exercise is to construct a minimum variance unbiased estimate for the proportion $\pi$ of a Normal population that is above a given value $c$ on the basis of an observed sample $x_1, x_2, \ldots, x_n$.

1. Recall that $\bar{X}$ and the unbiased estimate of variance $S^2$ are sufficient for $\mu$ and $\sigma^2$.

2. Consider the statistic $W$ defined by:

$$W = \begin{cases} 
0 & \text{if } x_1 \leq c \\
1 & \text{if } x_1 > c 
\end{cases}$$

Where $x_1$ is the first observation. It is easy to see that that $W$ is an unbiased estimate of $\pi$.

3. Use the two following facts to write out how to implement the “Rao-Blackwellisation” estimator based on $W$ and the sufficient statistics $\bar{X}$ and $S$.

(a) Suppose $U \sim \Gamma(\alpha, p_1)$ and $V \sim \Gamma(\alpha, p_2)$ are independent gamma random variables, then $\frac{U}{U+V}$ is a r.v. which is independent of $U+V$ and has a density $b(x)$ proportional to $x^{p_1-1}(1-x)^{p_2-1}$.

(b) The Beta distribution with parameters $\gamma$ and $\delta$ is the distribution defined by the density:

$$b(x|\gamma, \delta) = \frac{\beta(\gamma, \delta)}{\beta(\gamma + \delta)} x^{\gamma-1} (1-x)^{\delta-1}$$

where $\beta(\gamma, \delta) = \Gamma(\gamma)\Gamma(\delta)/\Gamma(\gamma + \delta)$ is a normalizing constant to make it a density.

**Question 1** Estimate the mean and the variance of Beta(0.5, 6.5) using the `betarnd` function.

**Question 2** Calculate $P(B \leq 0.02)$ with $B \sim Beta(0.5, 6.5)$ using the `betacdf` function.

The minimum variance unbiased estimate for $\pi$ is

$$E(W|\bar{X}, S) = P(X_1 > c|\bar{X}, S)$$

$$= P \left( \frac{X_1 - \bar{X}}{S} > \frac{c - \bar{X}}{S} |\bar{X}, S \right)$$

$$= P \left( \frac{X_1 - \bar{X}}{S} > \frac{c - \bar{X}}{S} \right) \quad \text{(Basu’s theorem)}$$

Consider the distribution of

$$Y = \frac{n}{n-1} \left( \frac{X_1 - \bar{X}}{S} \right)^2$$

1
Use the facts that $X_1 - \bar{X} \sim N(0, \frac{n-1}{n} \sigma^2)$, $(n - 1)S^2/\sigma^2 - \frac{m}{n-1}(X_1 - \bar{X})^2/\sigma^2 \sim \chi^2_{n-2}$ and they are independent of each other.

**Question 3** What is the distribution of $Y$?

Note that the $\chi^2_n$ distribution is a special case of the gamma distribution, with $\alpha = 1/2$ and $p = n/2$.

**Question 4** What is $E(W|\bar{X}, S)$?

Note that the distribution of $\frac{\sqrt{n-1}(X_1 - \bar{X})}{\sqrt{(n-1)S}}$ is symmetric.

**Question 5** Write a Matlab function to calculate the minimum variance unbiased estimate for $\pi$.

**Question 6** Obtain the estimate of $\pi$ for $c = 4.8$ obtained by this method applied to the sample

$$3.15, 2.92, 4.59, 6.48, 4.28, 4.81, 3.36, 3.20, 4.16, 6.48, 3.85, 3.72, 5.15, 2.67, 2.11$$

(For these numbers, $\bar{X} = 4.06, S^2 = 1.647$)