University of California, Santa Cruz
Department of Applied Mathematics and Statistics
Baskin School of Engineering
Classical and Bayesian Inference - AMS 132

Review and Example: Sampling distributions

1. Suppose that $X_1, ..., X_n$ form a random sample from the normal distribution with mean μ and variance σ^2, and Y is an independent random variable having the normal distribution with mean 0 and variance $4\sigma^2$.

   (a) Determine a function of $X_1, ..., X_n$ and Y that does not involve μ or σ^2 but has he t distribution with $n-1$ degrees of freedom.

   (b) If a sample of size 13 is considered, compute $P\left(\frac{Y}{2\sigma'} < 1.6\right)$, where $\sigma' = \sqrt{\frac{\sum_{i=1}^n (X_i - \overline{X}_n)^2}{n-1}}$.

   (c) If a sample of size 13 is considered, find values a and b, $a < b$ such that $P\left(a < \frac{Y}{2\sigma'} < b\right) = 0.94$.

2. Suppose that $X_1, ..., X_n$ are random variables describing the monthly income (in thousands of dollars) of people in the city of Santa Cruz. Assume that they form a random sample from the normal distribution with mean μ and variance σ^2.

   (a) Assume that both μ and σ^2 are unknown. Determine the smallest value of n such that the expected squared length of this interval will be less than $\sigma^2/2$.

   (b) If σ^2 is known, determine the smallest value of n such that the expected squared length of this interval will be less than $\sigma^2/2$.

   (c) Assume that both μ and σ^2 are unknown. From a sample of 25 person, the sample mean of their monthly income is 6.1 with a sample standard deviation of 1.2. Find a coefficient 0.9 confidence interval for the mean. The major has decided that a bonus will be given if the mean monthly income is less that 4 thousand dollars. What do you think the mayor will do based on this information?

   (d) Assume that σ^2 is known and equal to 1.5. Additionally, assume that a new person is considered in the sample. Find a coefficient 0.9 confidence interval for the monthly income of this new person.
   For this: find a pivot that has the standard normal distribution, and involves the monthly income of this new person and the other person in the sample. Then find random variables A and B such that $P(A < X_{n+1} < B) = 0.9$.

3. Suppose that $X_1, \ldots, X_7$ describe the age of 7 male students in a class and $Y_1, \ldots, Y_7$ describe the age of 7 female students in a class. In class, we assumed that $X_i \overset{iid}{\sim} N(\mu_1, \sigma^2)$, $Y_i \overset{iid}{\sim} N(\mu_2, \sigma^2)$, they are independent, and the variance is known an equal to 2. Now, lets assume that the variances are

unknown and there is uncertainty whether the same variance should be considered or not. For this, assume now that $X_1, \ldots, X_7$ describe the age of 7 male students in a class and $Y_1, \ldots, Y_7$ describe the age of 7 female students in a class. So, $X_i \overset{iid}{\sim} N(\mu_1, \sigma_1^2)$, $Y_i \overset{iid}{\sim} N(\mu_2, \sigma_2^2)$, and they are independent. It is of interest to find a 90 percent confidence interval for the ratio of σ_2^2 and σ_1^2. For this:

(a) Show that $Z_1 = \sum_{i=1}^n \left(\frac{X_i - \overline{X}_n}{\sigma_1}\right)^2 \sim \chi^2_{(n-1)}$, $Z_2 = \sum_{i=1}^n \left(\frac{Y_i - \overline{Y}_n}{\sigma_2}\right)^2 \sim \chi^2_{(n-1)}$, and they are independent. Give all the details on how you get this distribution.

(b) Find the random variable W such that $W\frac{\sigma_2^2}{\sigma_1^2}$ is a pivotal that has a Fisher distribution with $n-1$ and $n-1$ degrees of freedom. Give all the details on how you get this random variable and distribution.

Note: if X and Y are independent random variables such that $X \sim \chi^2_{m_1}$ and $Y \sim \chi^2_{m_2}$ then $\frac{X/m_1}{Y/m_2}$ follows a Fisher distribution with m_1 and m_2 degrees of freedom, denoted F_{m_1,m_2}.

(c) Consider $\gamma_1 = 0.05$ and $\gamma_2 = 0.95$, and compute $G^{-1}(\gamma_1)$ and $G^{-1}(\gamma_2)$, where G is the c.d.f. of a random variable that has a Fisher distribution with $n-1$ and $n-1$ degrees of freedom. You can use the table at the end of the book, page 862, or the R command, `qf(p, df1, df2` where `p` is the quantile, and `df1` and `df2` are the degrees of freedom.

(d) Find random variables A and B such that $P(A < \frac{\sigma_2^2}{\sigma_1^2} < B) = \gamma$.

(e) If $\sum_{i=1}^n (x_i - \overline{x}_n)^2 = 5.46$ and $\sum_{i=1}^n (y_i - \overline{y}_n)^2 = 11.05$, compute the γ percent confidence interval. What can you say about the equal variances assumption?