Homework 2

Solutions

1. Suppose $X_1, X_2, ..., X_{25}$ are i.i.d. $N(\mu, \sigma^2)$ where μ and σ are appropriately defined and assumed known.Construct a statistic that is a function of all of the available random variables($X_1, ..., X_n$) that is distributed

   a) $N(\mu, \frac{\sigma^2}{25})$

   b) $N(0, 1)$

   c) χ^2_{24}

   d) t_{24}

   Solution:

   a) $\bar{X} = \frac{1}{25}\sum_{i=1}^{25} X_i$

   b) $\frac{1}{5}\sum_{i=1}^{25} \frac{X_i-\mu}{\sigma}$

   c) $\frac{24}{\sigma^2}S^2 = \frac{1}{\sigma^2}\sum_{i=1}^{25}(X_i - \bar{X})^2$

   d) $\frac{\frac{1}{5}\sum_{i=1}^{25}\frac{X_i-\mu}{\sigma}}{\sqrt{\frac{24}{\sigma^2}S^2/24}} = \frac{\frac{1}{5}\sum_{i=1}^{25}(X_i-\mu)}{\sqrt{S^2}}$

2. Let $T \sim t_n$. Find the distribution of T^2 *Solution:*
   Since $T \sim t_n$, we know that by definition this means that we can write T as

$$T = \frac{Z}{\sqrt{Y/n}}$$

   Where $Z \sim N(0, 1)$ and $Y \sim \chi^2_n$ where Z and Y are independent. Therefore,

$$T^2 = \frac{Z^2}{Y/n}$$

   We know that since $Z \sim N(0, 1)$ $Z^2 \sim \chi^2_1$ (see problem 4 on quiz 1). Therefore $T^2 \sim F_{1,n}$

3. Let $F \sim F_{n_1,n_2}$ find the distribution of $\frac{1}{F}$ *Solution:* Since $F \sim F_{n_1,n_2}$, we know that by definition this means that we can write F as

$$F = \frac{W_1/n_1}{W_2/n_2}$$

Where $W_1 \sim \chi^2_{n_1}$ and $W_2 \sim \chi^2_{n_2}$ where W_1 and W_2 are independent. Therefore,

$$\frac{1}{F} = \frac{W_2/n_2}{W_1/n_1}$$

Therefore $\frac{1}{F} \sim F_{n_2,n_1}$

4. Suppose that you work for Coca-cola as a quality consultant in a bottling factory. One of the bottling machines seems to be dispersing less soda per bottle on average, so you take a sample of 25 bottles from the assembly line that the questionable machine works on and measure how much soda is in each bottle. For comparison, you also grab 16 bottles that are filled by a different machine. You assume that the amount that the bottles are filled by these machines is normally distributed with a mean value of 16 ounces and an unknown variance of $\sigma^2 > 0$. Let $X_1, ..., X_{25}$ be the amounts of soda in the 25 bottles (in ounces) from the machine that is under question and let $Y_1, ..., Y_{16}$ be the amounts of soda in the 16 bottles (again, in ounces) from other machine. You also assume that the amount a given bottle is filled is independent of the amount a different bottle is filled.

   a) Find the distribution of $U = \frac{5(\bar{X}-16)}{\sqrt{S_X^2}}$ where $\bar{X} = \frac{1}{25}\sum_{i=1}^{25} X_i$, and $S_X^2 = \frac{1}{24}\sum_{i=1}^{25}(X_i - \bar{X})^2$

   Solution:

$$\begin{aligned} U &= \frac{5(\bar{X}-16)}{\sqrt{S_X^2}} \\ &= \frac{\frac{5}{\sigma}(\bar{X}-16)}{\frac{1}{\sigma}\sqrt{24S_X^2/24}} \\ &= \frac{5\frac{\sum_{i=1}^{25}(\frac{X_i-16}{\sigma})}{25}}{\sqrt{\frac{24}{\sigma^2}S_X^2/24}} \\ \Rightarrow &\ U \sim t_{24} \end{aligned}$$

   Note:
   $\frac{X_i-16}{\sigma} \sim N(0,1)$ for all i.
   $\Rightarrow \frac{\sum_{i=1}^{25}(\frac{X_i-16}{\sigma})}{25} \sim N(0, \frac{1}{25})$
   $\Rightarrow 5\frac{\sum_{i=1}^{25}(\frac{X_i-16}{\sigma})}{25} \sim N(0,1)$
   Additionally, we know that
   $\frac{24}{\sigma^2}S_X^2 \sim \chi^2_{24}$
   And since $\bar{X}$ and S_X^2 are independent, $\frac{24}{\sigma^2}S_X^2$ and $\frac{5}{\sigma}(\bar{X}-16) = 5\frac{\sum_{i=1}^{25}(\frac{X_i-16}{\sigma})}{25}$ must also be independent.

   b) Find the distribution of $W = \frac{S_X^2}{S_Y^2}$, where $S_Y^2 = \frac{1}{15}\sum_{i=1}^{16}(Y_i - \bar{Y})^2$.

   Solution:

$$\begin{aligned}
W &= \frac{S_X^2}{S_Y^2} \\
&= \frac{\frac{1}{\sigma^2}\, 24S_X^2/24}{\frac{1}{\sigma^2}\, 15S_Y^2/15} \\
&= \frac{\frac{24}{\sigma^2}S_X^2/24}{\frac{15}{\sigma^2}S_Y^2/15} \\
&\Rightarrow W \sim F_{24,15}
\end{aligned}$$

5. Let $T \sim t_n$ where $n > 2$. By definition this means that we can write T as

$$T = \frac{Z}{\sqrt{Y/n}}$$

Where $Z \sim N(0,1)$ and $Y \sim \chi_n^2$ where Z and Y are independent.

a) Find $E[T]$

b) Find $V[T]$

Solution:

a)

$$\begin{aligned}
E[T] &= E[\frac{Z}{\sqrt{Y/n}}] \\
&= E[Z\frac{1}{\sqrt{Y/n}}] \\
&= E[Z]E[\frac{1}{\sqrt{Y/n}}] \leftarrow \text{ since } Z \text{ and } Y \text{ are independent} \\
&= 0 * E[\frac{1}{\sqrt{Y/n}}] \leftarrow \text{ Since } Z \sim N(0,1) \\
&= 0
\end{aligned}$$

b)

$$\begin{aligned}
V[T] &= E[T^2] - E^2[T] \\
&= E[T^2] \leftarrow \text{ since, in part a, we showed that } E[T] = 0 \\
&= E[\frac{Z^2}{Y/n}] \\
&= E[Z^2\frac{1}{Y/n}] \\
&= E[Z^2]E[\frac{1}{Y/n}] \leftarrow \text{ since } Z \text{ and } Y \text{ are independent} \\
&= nE[Z^2]E[\frac{1}{Y}]
\end{aligned}$$

$$
\begin{aligned}
&= n(\frac{1}{2})(2)(2^{-1})(\frac{\Gamma(\frac{n}{2}-1)}{\Gamma(\frac{n}{2})}) \text{ *see below} \\
&= \frac{n}{n-2}
\end{aligned}
$$

$^*E[Z^2] = (\frac{1}{2})(2)$ because $Z^2 \sim \chi_1^2 \equiv \Gamma(\frac{1}{2}, 2)$
$E[\frac{1}{Y}] = E[Y^{-1}] = 2^{-1}\frac{\Gamma(\frac{n}{2}-1)}{\Gamma(\frac{n}{2})}$ because $Y \sim \chi_n^2 \equiv \Gamma(\frac{n}{2}, 2)$

6. Problem 7.9 from the book (p 364)
   Solution:
   a. $P(|\overline{Y} - \mu| \leq .3) = P(-1.2 \leq Z \leq 1.2) = .7698$.
   b. $P(|\overline{Y} - \mu| \leq .3) = P(-.3\sqrt{n} \leq Z \leq .3\sqrt{n}) = 1 - 2P(Z > .3\sqrt{n})$. For n = 25, 36, 69, and 64, the probabilities are (respectively) .8664, .9284, .9642, and .9836.
   c. The probabilities increase with n, which is intuitive since the variance of $\overline{Y}$ decreases with n.
   d. Yes, these results are consistent since the probability was less than .95 for values of n less than 43.

7. Problem 7.10 from the book (p 364)
   Solution:
   a. $P(|\overline{Y} - \mu| \leq .3) = P(-.15\sqrt{n} \leq Z \leq .15\sqrt{n}) = 1 - 2P(Z > .15\sqrt{n})$. For $n = 9$, the probability is .3472 (a smaller value).
   b. For $n = 25$: $P(|\overline{Y} - \mu| \leq .3) = 1 - 2P(Z > .75) = .5468$
   For $n = 36$: $P(|\overline{Y} - \mu| \leq .3) = 1 - 2P(Z > .9) = .6318$
   For $n = 49$: $P(|\overline{Y} - \mu| \leq .3) = 1 - 2P(Z > 1.05) = .7062$
   For $n = 64$: $P(|\overline{Y} - \mu| \leq .3) = 1 - 2P(Z > 1.2) = .7698$
   c. The probabilities increase with n.
   d. The probabilities are smaller with a larger standard deviation (more diffuse density).

8. Problem 7.15 (a and b only) from the book (p. 365)
   Solution:
   We know that $X_1, ..., X_m$ are i.i.d. $N(\mu_1, \sigma_1^2)$. Therefore $\bar{X} \sim N(\mu_1, \frac{\sigma_1^2}{m})$. Similarly, we can deduce that $\bar{Y} \sim N(\mu_2, \frac{\sigma_2^2}{n})$. This implies that $-\bar{Y} \sim N(-\mu_2, \frac{\sigma_2^2}{n})$.
   Therefore $\bar{X} - \bar{Y} = \bar{X} + -\bar{Y} \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n})$, since $X_1, ..., X_m, Y_1, ..., Y_n$ are all independent.

   Therefore

   a) $E[\bar{X} - \bar{Y}] = \mu_1 - \mu_2$
   b) $E[\bar{X} - \bar{Y}] = \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}$

9. Problem 7.20 from the book (p.366) *Solution:*

a. Using the fact that the chi–square distribution is a special case of the gamma distribution, $E(U) = \nu$, $V(U) = 2\nu$.

b. Using Theorem 7.3 and the result from part a:

$$E(S^2) = \tfrac{\sigma^2}{n-1} E(\tfrac{n-1}{\sigma^2} S^2) = \tfrac{\sigma^2}{n-1}(n-1) = \sigma^2.$$

$$V(S^2) = \left(\tfrac{\sigma^2}{n-1}\right)^2 V(\tfrac{n-1}{\sigma^2} S^2) = \left(\tfrac{\sigma^2}{n-1}\right)^2 [2(n-1)] = 2\sigma^4/(n-1).$$

Challenge Question:

Let $F \sim F_{n_1,n_2}$ where $n_1 > 4$ and $n_2 > 4$. Find both the mean and the variance of F.