Solutions to Assignment #5

1. Let X denote a random variable having a normal(μ, σ^2) distribution. Define

$$Z = \frac{X - \mu}{\sigma}.$$

Compute the mgf of Z and use it to deduce the distribution of Z.

***Solution**:* Compute

$$\begin{aligned} M_Z(t) &= E(e^{tZ}) \\ &= E\left(e^{(X-\mu)\frac{t}{\sigma}}\right) \\ &= E\left(e^{-\mu t/\sigma} e^{X\left(\frac{t}{\sigma}\right)}\right) \\ &= e^{-\mu t/\sigma} E\left(e^{X\left(\frac{t}{\sigma}\right)}\right) \\ &= e^{-\mu t/\sigma} M_X\left(\frac{t}{\sigma}\right), \end{aligned}$$

where

$$M_X\left(\frac{t}{\sigma}\right) = e^{\mu t/\sigma + \sigma^2 (t/\sigma)^2/2} = e^{\mu t/\sigma} \cdot e^{t^2/2}.$$

It then follows that

$$M_Z(t) = e^{t^2/2}, \quad \text{for all } \ t \in \mathbb{R},$$

which is the mgf of a normal$(0, 1)$ distribution. Consequently, $Z \sim$ normal$(0, 1)$. □

2. Let $X_1, X_2, \ldots, X_n$ denote a random sample from a normal(μ, σ^2) distribution. Define

$$Z_n = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}},$$

where $\overline{X}_n$ denotes the sample mean. Compute the mgf of Z_n and use it to deduce the distribution of Z_n.

Solution: We have seen in class that $\overline{X}_n$ has a normal$(\mu, \sigma^2/n)$ distribution. Consequently, by the result of the previous problem,

$$\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim \text{normal}(0,1).$$

Thus, Z_n has a normal$(0,1)$ distribution. □

3. Let $Z \sim$ normal$(0,1)$ and define $X = \mu + \sigma Z$. Compute the mgf of X and use it to deduce the distribution of X.

Solution: Compute

$$\begin{aligned} M_X(t) &= E(e^{\mu t + \sigma t Z}) \\ &= e^{\mu t}\, E(e^{\sigma t Z}) \\ &= e^{\mu t}\, M_Z(\sigma t), \end{aligned}$$

where

$$M_Z(\sigma t) = e^{(\sigma t)^2/2} = e^{\sigma^2 t^2/2},$$

so that

$$M_X(t) = e^{\mu t + \sigma^2 t^2/2}, \quad \text{for all } \; t \in \mathbb{R},$$

which is the mgf of a normal(μ, σ^2) distribution. Consequently, X has a normal(μ, σ^2) distribution. □

4. Let $\beta > 0$ and $X_1, X_2, \ldots, X_n$ be a random sample from an exponential(β) distribution.

   Define $Y_n = \dfrac{2n\overline{X}_n}{\beta}$, where $\overline{X}_n$ is the sample mean

   (a) Determine the distribution of Y_n.

   Solution: We saw in part (c) of problem 2 in Assignment #1 that the mgf of Y_n is

$$M_{Y_n}(t) = \left(\frac{1}{1-2t}\right)^n = \left(\frac{1}{1-2t}\right)^{2n/2} \quad \text{for } \; t < \frac{1}{2},$$

   which is the mgf for a $\chi^2(2n)$ distribution. Thus, $Y_n \sim \chi^2(2n)$. □

(b) For $n = 10$, find values of c and d so that

$$\mathrm{P}\left(c < \frac{2n\overline{X}_n}{\beta} < d\right) \doteq 0.95.$$

Use this result to give a 95% confidence interval for β based on the sample mean.

Solution: By the result of the previous part,

$$\mathrm{P}\left(c < \frac{2n\overline{X}_n}{\beta} < d\right) = \mathrm{P}(c < Y_n < d),$$

where Y_n has a χ^2 distribution with $2n$ degrees of freedom, or 20 degrees of freedom in this case. Let F_{Y_n} denote the cdf of Y_n. Then,

$$\mathrm{P}(c < Y_n < d) = F_{Y_n}(d) - F_{Y_n}(c).$$

to get $\mathrm{P}(c < Y_n < d) = 0.95$ we may choose c so that $F_{Y_n}(c) = 0.025$ and d so that $F_{Y_n}(d) = 0.975$. Thus,

$$c = F_{Y_n}^{-1}(0.025) \quad \text{and} \quad d = F_{Y_n}^{-1}(0.975).$$

Using R or MS Excel we obtain values of c and d. In R use the `qchisq` function to get

$$c \approx \texttt{qchisq}(0.025, \texttt{df} = 20) \approx 9.59$$

and

$$d \approx \texttt{qchisq}(0.975, \texttt{df} = 20) \approx 34.17.$$

In MS Excel, the function `CHIINV` returns the inverse of the right–tail probability for the χ^2 distribution; in other words,

$$\texttt{CHIINV}(\texttt{probability, df}) = 1 - F_{Y_n}^{-1}(1 - \texttt{probability}).$$

we therefore have that the 95% confidence interval for β based on the sample mean can be obtained from

$$9.59 < \frac{20\overline{X}_n}{\beta} < 34.17,$$

from which we get

$$\frac{20}{34.17}\overline{X}_n < \beta < \frac{20}{9.59}\overline{X}_n.$$

Thus, the 95% confidence interval for β is

$$\left(0.59\overline{X}_n, 2.09\overline{X}_n\right).$$

□

5. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Poisson distribution with mean λ. Thus, $Y = \sum_{i=1}^{n} X_i$ has a Poisson distribution with mean $n\lambda$. Moreover, by the Central Limit Theorem, $\overline{X} = Y/n$ has, approximately, a normal$(\lambda, \lambda/n)$ distribution for large n.

   (a) Give the distribution of approximate distribution of

   $$\frac{Y/n - \lambda}{\sqrt{\lambda}/\sqrt{n}}$$

   for large values of n.

   Solution: Since, Y/n is the sample mean, with expected value λ and variance λ/n, the central limit theorem implies that

   $$\frac{Y/n - \lambda}{\sqrt{\lambda}/\sqrt{n}} \sim \text{normal}(0, 1)$$

   for large values of n. □

   (b) By the weak law of large numbers $|Y/n - \lambda|$ is very close to 0 for large values of n with a very high probability (i.e., probability very close to 1). Use this fact to obtain the approximation

   $$\sqrt{Y/n} \approx \sqrt{\lambda} + \frac{1}{2\sqrt{\lambda}}(Y/n - \lambda)$$

   for large values of n and very high probability.

   Solution: Using the first order approximation around $t = \lambda$ for the function $g(t) = \sqrt{t}$; namely,

   $$g(t) \approx g(\lambda) + g'(\lambda)(t - \lambda) \quad \text{for } t \text{ close to } \lambda,$$

   we obtain that

   $$\sqrt{Y/n} \approx \sqrt{\lambda} + \frac{1}{2\sqrt{\lambda}}(Y/n - \lambda) \quad \text{for large } n.$$

   □

(c) Prove that, for large values of n,

$$\mathrm{P}\left(2\sqrt{n}\left(\sqrt{Y/n}-\sqrt{\lambda}\right)\leqslant z\right)\approx \mathrm{P}(Z\leqslant z)\quad\text{for all}\quad z\in\mathbb{R}.$$

Solution: From the result of the previous part we have that

$$2\sqrt{n}(\sqrt{Y/n}-\sqrt{\lambda})\approx\frac{Y/n-\lambda}{\sqrt{\lambda}/\sqrt{n}}$$

so that, by the result from part (a), approximately,

$$2\sqrt{n}(\sqrt{Y/n}-\sqrt{\lambda})\sim \text{normal}(0,1)$$

for large values of n. Thus,

$$\mathrm{P}\left(2\sqrt{n}\left|\sqrt{Y/n}-\sqrt{\lambda}\right|<z\right)\approx \mathrm{P}(|Z|<z)\quad\text{for}\quad z>0$$

□

(d) Explain how you would use the result of part (c) to obtain a confidence interval estimate for the parameter λ.

Solution: Choosing $z_{\alpha/2}$ for that

$$\mathrm{P}(|Z|<z_{\alpha/2})=1-\alpha,$$

we obtain the $100(1-\alpha)\%$ confidence interval for λ as follows: First, compute that the approximate $100(1-\alpha)\%$ confidence interval for $\sqrt{\lambda}$

$$\sqrt{\frac{Y}{n}}-\frac{z_{\alpha/2}}{2\sqrt{n}}<\sqrt{\lambda}<\sqrt{\frac{Y}{n}}+\frac{z_{\alpha/2}}{2\sqrt{n}}.$$

We can then square all terms in the inequality to obtain an approximate $100(1-\alpha)\%$ confidence interval for λ:

$$\left(\sqrt{\frac{Y}{n}}-\frac{z_{\alpha/2}}{2\sqrt{n}}\right)^2<\lambda<\left(\sqrt{\frac{Y}{n}}+\frac{z_{\alpha/2}}{2\sqrt{n}}\right)^2.$$

□