Solutions to Exam #1

1. Define the following terms:

   (a) Random sample

   Answer: A random sample of size n from a given distribution is a set of independent random variables, $X_1, X_2, \ldots, X_n$, which have the same distribution as that from which sampling is being done. □

   (b) Statistic

   Answer: A statistic is a quantity computed from the values of a random sample; thus, a statistic random variable defined in terms of a random sample, $X_1, X_2, \ldots, X_n$. □

   (c) Sampling distribution

   Answer: The distribution of a statistic is called the sampling distribution of the statistic. □

   (d) Unbiased estimator

   Answer: Let T_n denote a statistic based on a random sample of size n from a distribution with parameter θ. T_n is said to be an unbiased estimator for θ if

$$E(T_n) = \theta.$$

□

   (e) Consistent estimator

   Answer: Let T_n denote a statistic based on a random sample of size n from a distribution with parameter θ. T_n is said to be a consistent estimator for θ if T_n converges to θ in probability as $n \to \infty$; that is, for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathrm{P}(|T_n - \theta| < \varepsilon) = 1.$$

□

2. Let X and Y be random variables with $X \sim \chi^2(1)$ and $Y \sim \chi^2(n)$ for $n > 1$, and define
$$W = Y - X.$$
Assuming that X and W are independent, determine the distribution of W.

Suggestion: Write $Y = X + W$ and compute the mgf of Y in terms of the mgfs of X and W.

Solution: Assume that X and W are independent and write
$$Y = W + X.$$
Then,
$$M_Y(t) = M_W(t) \cdot M_X(t),$$
by the independence assumption. Consequently,
$$\begin{aligned} M_W(t) &= \frac{M_Y(t)}{M_X(t)} \\ &= \frac{\left(\dfrac{1}{1-2t}\right)^{n/2}}{\left(\dfrac{1}{1-2t}\right)^{1/2}} \\ &= \left(\frac{1}{1-2t}\right)^{(n-1)/2}, \end{aligned}$$
which is the mgf for a $\chi^2(n-1)$ random variable. Therefore, W has a χ^2 distribution with $n-1$ degrees of freedom. □

3. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Poisson(λ) distribution and define the statistic $Y = \sum_{i=1}^{n} X_i$.

   (a) Derive the sampling distribution for Y. Justify your answer.

   Solution: Compute the mgf of Y, $M_Y(t) = E(e^{tY})$, to get that
$$M_Y(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdots M_{X_n}(t),$$

where we have used the independence assumption. Thus, since the random variables $X_1, X_2, \ldots, X_n$ are identically distributed, it follows that

$$M_Y(t) = \left(e^{\lambda(e^t-1)}\right)^n = e^{n\lambda(e^t-1)},$$

which is the mgf of a Poisson($n\lambda$) random variable. It follows that Y is a Poisson random variable with parameter $n\lambda$. □

(b) Find a value of c for that $T = cY$ is an unbiased estimator for λ. Justify your answer.

Solution: Since $Y \sim$ Poisson($n\lambda$), by part (a), it follows that $E(Y) = n\lambda$. Consequently,

$$E\left(\frac{1}{n}Y\right) = \lambda,$$

which shows that $T = \dfrac{1}{n}Y$ is an unbiased estimator for λ. □

4. Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal(μ, σ^2) distribution and define the statistic

$$T_n = \sum_{i=1}^{n}(X_i - \overline{X}_n)^2,$$

where $\overline{X}_n$ denotes the sample mean. We will show later in this course that $\dfrac{1}{\sigma^2}T_n$ has a χ^2 distribution with $n-1$ degrees of freedom.

(a) Explain how you would use knowledge of the distribution of $\dfrac{1}{\sigma^2}T_n$ to obtain a $100(1-\alpha)\%$ confidence interval for the variance σ^2 of a normal(μ, σ^2) distribution based on a random sample of size n from that distribution.

Solution: Let $Y = \dfrac{1}{\sigma^2}T_n$. Given that $Y \sim \chi^2(n-1)$, where n is known, we can find c and d so that

$$F_Y(c) = \frac{\alpha}{2} \quad \text{and} \quad F_Y(d) = 1 - \frac{\alpha}{2}.$$

It then follows that

$$\mathrm{P}(c < Y < d) = F_Y(d) - F_Y(c) = 1 - \alpha,$$

where we have used the fact that Y is a continuous random variable. It then follows that

$$\mathrm{P}\left(c < \frac{1}{\sigma^2}T_n < d\right) = 1-\alpha,$$

from which we get that

$$\mathrm{P}\left(\frac{1}{d} < \frac{\sigma^2}{T_n} < \frac{1}{c}\right) = 1-\alpha,$$

or

$$\mathrm{P}\left(\frac{1}{d}T_n < \sigma^2 < \frac{1}{c}T_n\right) = 1-\alpha.$$

Thus,

$$\left(\frac{1}{d}T_n, \frac{1}{c}T_n\right)$$

is a $100(1-\alpha)\%$ confidence interval for the variance σ^2. □

(b) Give a 95% confidence interval for the variance of a normal(μ, σ^2) distribution based on the statistic T_n, where the sample size, n, is 17.

Solution: Here, $\alpha = 0.05$ and $Y \sim \chi^2(16)$. Therefore,

$$c = F_Y^{-1}(0.025) = 6.91 \quad \text{and} \quad d = F_Y^{-1}(0.975) = 28.8.$$

we then have that a 95% confidence interval for the variance in this case is

$$\left(\frac{1}{28.8}T_n, \frac{1}{6.91}T_n\right). \tag{1}$$

□

(c) Assume that the counts of popcorn kernels in a 1/4 cup follow a normal distribution with parameters μ and σ^2, which are unknown. Seventeen students in this class measured a 1/4 cup of kernels and counted the kernels in the the container. The value of T_n for this particular sample of size $n = 17$ is about $21,900$. Use this information to provide a 95% confidence interval for the variance, σ^2. Give an interpretation of your result.

Solution: Using the result of the previous part in (1) we get that

$$\left(\frac{21,900}{28.8}, \frac{21,900}{6.91}\right),$$

or about

$$(760, 3560),$$

is a 95% confidence interval for the variance of the number kernels in 1/4 cup of popcorn. This interval might or might not contain the true variance, but we are confident that, on average, 95% of the intervals given by the formula in (1) computed from data in samples of size 17 will contain the true variance. □