Statistics I, Fall 2012
Homework 09

Instructor: Ling-Chieh Kung
Department of Information Management
National Taiwan University

1. (10 points) Let $\{X_i\}_{i=1,\dots,n}$ be a random sample where $X_i \sim \text{ND}(\mu, \sigma)$ and X_i and X_j are independent for all $i \neq j$. Prove that

$$\frac{\sum_{i=1}^{n}(X_i - \mu)^2}{\sigma^2} \sim \text{Chi}(n).$$

Hint. Use (1) the lemma for the square of a standard normal random variable proved in the lecture and (2) the fact that the moment generating function of a sum of independent random variables is the product of the moment generating functions of these random variables.

2. (10 points) Recall that in the proof of[1]

$$\frac{(n-1)S^2}{\sigma^2} \sim \text{Chi}(n-1) \quad \text{if } X_i \sim \text{ND}(\mu, \sigma) \text{ and } X_i \perp X_j \forall i \neq j,$$

we have constructed a matrix A and a column vector $Y = AZ$, where A is an orthogonal matrix and $Z = \begin{bmatrix} Z_1 & Z_2 & \cdots Z_n \end{bmatrix}^T$. In the lecture we have proved that $Y_2 \sim \text{ND}(0, 1)$ and $Y_3 \sim \text{ND}(0, 1)$. Replicate the proof and prove that $Y_n \sim \text{ND}(0, 1)$.

3. (10 points) Prove that

$$\sum_{i=1}^{n} Z_i^2 - n\overline{Z}^2 = \sum_{i=1}^{n}(Z_i - \overline{Z})^2,$$

where Z_is are random variables and $\overline{Z}$ is the mean of Z_is.

Hint. Go from the right-hand side may be easier.

4. (10 points) When we proved the central limit theorem in the lecture, we used a fact on the convergence of distribution. The fact is stated below. Let $\{X_n\}_{n=1,2,\dots}$ be a sequence of random variables and X be another random variable. Let $m_n(t)$ and $m(t)$ be the moment generating functions of X_n and X, respectively. Let $F_n(x)$ and $F(x)$ be the cumulative distribution functions of X_n and X, respectively. Then the fact is

$$\lim_{n\to\infty} m_n(t) = m(t) \quad \Rightarrow \quad \lim_{n\to\infty} F_n(x) = F(x),$$

i.e., the distributions of X_ns converge to the distribution of X. In this case, we say the limiting distribution of X_ns is the distribution of X. In this problem, we will demonstrate this fact by giving you an example.

   (a) (2 points) Let $\lambda_n = 1 - \frac{1}{n}$ be the rate of the exponential random variable X_n. Write down the moment generating function $m_n(t)$ and cumulative distribution function $F_n(x)$ of X_n, respectively. Your answer should not contain λ_n.

   (b) (3 points) Find $\lim_{n\to\infty} m_n(t)$.

   (c) (3 points) Find $\lim_{n\to\infty} F_n(x)$.

   (d) (2 points) Use your results in Parts (b) and (c) to determine the limiting distribution of X_ns as $n \to \infty$.

5. (25 points; 5 points each) In a population, 30% of entities are labeled as 1 and 70% as 0. Suppose we sample with replacement.

[1] $X_i \perp X_j$ means X_i and X_j are independent.

(a) Let X_1 be a randomly selected entity. What is the distribution of X_1? What is the probability that in the sample (with sample size 1) the number of 1s is more than the number of 0s?

(b) Let X_2 be another randomly selected entity. What is the distribution of $\hat{p} = \frac{X_1+X_2}{2}$? What is the probability that in the sample (with sample size 2) the number of 1s is more than the number of 0s?

(c) Let X_3 be another randomly selected entity. What is the distribution of $\hat{p} = \frac{X_1+X_2+X_3}{3}$? What is the probability that in the sample (with sample size 3) the number of 1s is more than the number of 0s?

(d) Let $\{X_i\}_{i=1,\ldots 50}$ be a random sample. What is the distribution of $\hat{p} = \frac{1}{50}\sum_{i=1}^{50} X_i$? What is the probability that in the sample (with sample size 50) the number of 1s is more than the number of 0s?

(e) Does increasing the sample size always decrease the probability of observing more 1s than 0s? Briefly explain why.

6. (15 points; modified from Problem 7.46 in the textbook) According to the US Bureau of Labor Statistics, 48% of all adults are women. Among all the adults, 25% of women and 20% of men have some volunteering experiences.

(a) (5 points) If we randomly sample 150 adult women, what is the probability of getting 35 or more people who volunteer?

(b) (10 points) If we randomly sample 300 adults, what is the probability that the sample proportion of people who volunteer is between 20% and 25%?

7. (20 points) Suppose you sample $n = 15$ values from a normal population with population variance $\sigma^2 = 10$. Let S^2 be the sample variance.

(a) (5 points) What is the probability that $\frac{(n-1)S^2}{\sigma^2} = 1.4S^2$ is above 10?

(b) (5 points) What is the probability that $\frac{(n-1)S^2}{\sigma^2} = 1.4S^2$ is between 6 and 14?

(c) (10 points) What is the probability that S^2 is between 6 and 14?