COLUMBIA UNIVERSITY

DEPARTMENT OF BIOSTATISTICS

P8109 – STATISTICAL INFERENCE

Exercise Sheet 1

Date due: MON Feb 10, 2020.

No late assignments will be accepted.

Question 1

Calculate $\Pr\left\{\sum_{i=1}^{6} Z_i^2 \le 6\right\}$ where Z_i's are i.i.d. $N(0,\ 1)$ r.v.'s.

Question 2

If X is a random variable that has an F-distribution with ν_1 numerator and ν_2 denominator degrees of freedom, show that $Y = 1/X$ has an F-distribution with ν_2 numerator and ν_1 denominator degrees of freedom

---

Question 3

Let

$$T = \frac{Z}{\sqrt{W/\nu}},$$

where $Z \sim N(0,\ 1)$ and W is an independent χ^2_ν -distribution.

(a) Write down the distribution of T.

(b) Using the result that

$$\mathcal{E}W^{\alpha} = \frac{\Gamma\left(\frac{\nu}{2} + \alpha\right)}{\Gamma\left(\frac{\nu}{2}\right)} 2^{\alpha} \quad \text{for} \quad \nu > -2\alpha,$$

show that:

(i) $\mathcal{E}T = 0$ for $\nu > 1$;

(ii) $\operatorname{var} T = \nu / (\nu - 2)$ if $\nu > 2$.

Question 4

The following is a proof of the Central Limit Theorem for the Poisson distribution, i.e. if $X \sim \text{Poisson}(\lambda)$, then as $\lambda \to \infty$, $(X - \lambda)/\sqrt{\lambda} \xrightarrow{D} Z \sim N(0,1)$:

(a) Define $X^* = (X - \lambda)/\sqrt{\lambda}$. Show that:

$$\mathcal{E}e^{tX^*} = e^{-t\sqrt{\lambda}}\mathcal{E}e^{tX/\sqrt{\lambda}}.$$

(b) Show that:

$$\mathcal{E}e^{tX^*} = e^{-t\sqrt{\lambda}-\lambda}\sum_{x=0}^{\infty}\frac{\left(\lambda e^{t/\sqrt{\lambda}}\right)^x}{x!}.$$

(c) By using the fact that $\sum_{i=0}^{\infty} a^i / i! \equiv e^a$, show that

$$\mathcal{E}e^{tX^*} = \exp\left\{-t\sqrt{\lambda} - \lambda + \lambda e^{t/\sqrt{\lambda}}\right\}.$$

(d) Hence show that

$$\mathcal{E}e^{tX^*} \to e^{t^2/2} \quad \text{as } \lambda \to \infty.$$

Question 5

Prove that if $X \sim \text{binomial}(n, p)$, then

$$\frac{X - np}{\sqrt{np(1-p)}} \xrightarrow{D} Z \sim N(0,1).$$

Hint: You may want to use the fact that X can be written as the sum of n i.i.d. Bernoulli r.v.'s, i.e. $X = Z_1 + Z_2 + ... + Z_n$, *where* $Z_i = 1$ *with prob. p and* $Z_i = 0$ *with prob. 1-p.*

Question 6

Let $Y_1, ..., Y_5$ be a random sample of size 5 from a normal population with mean 0 and variance 1. Let $\bar{Y} = (\sum_{i=1}^5 Y_i)/5$ and Y_6 be another independent observation from the same population. What is the distribution of:

(a) $\bar{Y}$;

(b) $U = \sum_{i=1}^5 Y_i^2$;

(c) $V = \sum_{i=1}^5 (Y_i - \bar{Y})^2$;

(d) $W = \sum_{i=1}^5 (Y_i - \bar{Y})^2 + Y_6^2$.

Question 7

The coefficient of variation (*CV*) for a random sample $Y_1, ..., Y_n$ is defined by

$$CV = \frac{S}{\bar{Y}},$$

where *S* is the sample standard deviation and $\bar{Y}$ is the sample mean. *CV* measures the amount of variation as a proportion of the sample mean. Suppose each $Y_i \sim N(0, \sigma^2)$ for $i = 1, ..., 10$.

(a) By using $F_{1,\nu} \equiv t_\nu^2$, find the distribution of $10\bar{Y}^2 / S^2$ in terms of the *F*-distribution.

(b) Find the distribution of $S^2 / (10\bar{Y}^2)$.

(c) Find the number *c* such that

$$\Pr\left\{-c \le \frac{S}{\bar{Y}} \le c\right\} = .95 .$$

Question 8

Consider a sequence of i.i.d. r.v.'s $Z_1, ..., Z_n$ where each $Z_i \sim N(0, 1)$.

(a) Write down the relationship between the Z_i 's and *W*, where $W \sim \chi_{n-1}^2$

(b) Now consider another sequence of i.i.d. r.v.'s $X_1, ..., X_n$ where each $X_i \sim N(\mu, \sigma^2)$ and let

$$S^2 = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1}$$

(i)Write down the relationship between S^2 and *W*.

(ii)Use the Central Limit Theorem to show that

$$\frac{S^2 - \sigma^2}{\sigma^2\sqrt{2/(n-1)}} \xrightarrow{D} Z_i \sim N(0, 1).$$

Question 9

Verify if each of the following distributions belongs to the exponential family:

(a) binomial(n, p);

(b) gamma(α, β).

Question 10

(a) Consider a r.v. *X* with expectation μ and let $h(X)$ be a function of *X* with continuous second derivative. Show that

$$\mathcal{E}h(X) \approx h(\mu) + \frac{1}{2}h''(\mu)\operatorname{var}(X).$$

(b) The probability of success *P* of a new medical procedure is estimated to be $P_s = .368$ based on $n = 378$ observations. Ignoring second derivatives in the formula in (a) above, find estimates for the expectation and variance of the odds of success:

$$\widehat{O} = \frac{P_s}{1 - P_s}.$$

ANSWERS

Q1 .57681**Q7** $c = 49.04$ **Q10** (b) .58, 3.86×10^{-3}

Dr. P Gorroochurn: 01/30/2020