Department of Mathematics
MTL 601 (Probability and Statistics)
Tutorial Sheet No. 6

1. If random samples of size three and drawn without replacement from the population consisting of four numbers $4, 5, 5, 7$. Find sample mean X for each sample and make sampling distribution of X. Calculate the mean and standard deviation of this sampling distribution. Compare your calculations with population parameters.

2. Assume that a school district has 10,000 6th graders. In this district, the average weight of a 6th grader is 80 pounds, with a standard deviation of 20 pounds. Suppose you draw a random sample of 50 students. What is the probability that the average weight of a sampled student will be less than 75 pounds?

3. Calculate the probability that, for a random sample of 5 values taken from a $N(100, 252)$ population
   (a) $\overline{X}$ will be between 80 and 120
   (b) S will exceed 41.7

4. Independent random samples of size n_1 and n_2 are taken from the normal populations $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ respectively
   (a) Write down the sampling distributions of $\overline{X_1}$ and $\overline{X_2}$ and hence determine the sampling distribution of $\overline{X_1} - \overline{X_2}$, the difference between the sample means.
   (b) Now assume that $\sigma_1^2 = \sigma_2^2 = \sigma^2$,
      (i) Express the sampling distribution of $\overline{X_1} - \overline{X_2}$ in standard normal form.
      (ii) State the sampling distribution of $\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{\sigma^2}$.
      (iii) Using the $N(0, 1)$ distribution from (a) and the χ^2 distribution from t distribution to find the sampling distribution of $\overline{X_1} - \overline{X_2}$ when σ^2 is unknown.

5. Prove that $\overline{X}$, the mean of a random sample of size n from a distribution that is $N(\theta, \sigma^2)$, $-\infty < \theta < \infty$ is an efficient estimator of θ for every known $\sigma^2 > 0$.

6. Assuming population to be $\mathcal{N}\ (\mu, \sigma^2)$, show that sample variance is a consistent estimator for population variance σ^2.

7. Let X and Y be two independent random variables. Show that $X + Y$ is normally distributed if and only if both X and Y are normal.

8. Let $X_i, i = 1, 2, \ldots, n$ be a random sample of size n drawn from a population of size N with population distribution given by
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty.$$
   Show that $\sum_{i \neq j} Cov(X_i, X_j) = n(n-1)\left(\frac{-\sigma^2}{N-1}\right)$.

9. Let (X, Y) be a random vector of continuous type with joint PDF $f(x, y)$. Define $Z = X+Y$, $U = X-Y$, $V = XY$, $W = \frac{X}{Y}$. Then, prove that the PDFs of Z, U, V, and W are, respectively, given by
$$f_Z(z) = \int_{-\infty}^{\infty} f(x, z-x)dx$$
$$f_U(u) = \int_{-\infty}^{\infty} f(u+y, y)dy$$
$$f_V(v) = \int_{-\infty}^{\infty} f(x, v/x)\frac{1}{|x|}dx$$
$$f_W(w) = \int_{-\infty}^{\infty} f(xw, x)|x|dx.$$

10. Let X and Y be two independent $N(0,1)$ random variables. Show that $X+Y$ and $X-Y$ are independent.

11. If X and Y are independent random variables with the same distribution and $X+Y, X-Y$ are independent. Show that all random variables $X, Y, X+Y, X-Y$ are normally distributed.

12. If X and Y are independent exponential random variable with parameter λ, then show that $\frac{X}{X+Y}$ follows uniform distribution on $(0,1)$.

13. If X_1 and X_2 are independent random variables such that $X_1 \sim Exp(\lambda)$ and $X_2 \sim Exp(\lambda)$, show that $Z = \min\{X_1, X_2\}$ follows $Exp(2\lambda)$. Hence, generalize this result for n independent exponential random variables.

14. Let $Z \sim N(0,1)$ and $X \sim \chi^2_n$. Prove that $Y = \frac{Z}{\sqrt{\frac{X}{n}}} \sim t(n)$.

15. Let $X \sim \chi^2_{\nu_1}$ and $Y \sim \chi^2_{\nu_2}$. Prove that $Z = \frac{X/\nu_1}{Y/\nu_2} \sim F(\nu_1, \nu_2)$.

16. Let $X_i \sim U(0,1),\ i = 1, 2, \ldots, n$ such that X_i are independent. Then, find the distribution of $-\sum_{i=1}^{n} \log(X_i)$.

17. Let $X_1, X_2, \ldots, X_n$ be a random sample from a Bernoulli distribution with the parameter $p,\ 0 < p < 1$. Prove that the sampling distribution of the sample variance S^2 is

$$P\left(S^2 = \frac{i(n-i)}{n(n-1)}\right) = {}^nC_i p^i (1-p)^{n-i} + {}^nC_{n-i} p^{n-i} (1-p)^i$$

where $i = 0, 1, 2, \ldots, [\frac{n}{2}]$

18. Let $X_1, X_2 \ldots, X_n$ be a random sample form a distribution function F, and let $F_n^*(x)$ be the empirical distribution function of the random sample. Prove that $Cov(F_n^*(x), F_n^*(y)) = \frac{1}{n} F(x)(1 - F(y))$.

19. Find $Cov(\overline{X}, S^2)$ in terms of population moments. Under what conditions is $Cov(\overline{X}, S^2) = 0$? Discuss this result when the population is normal distributed.

20. Find the probability that the maximum of a random sample of size n from a population exceeds the population median.

21. Consider repeated observation on a $m-$dimensional random variable with mean $E(X_i) = \mu, i = 1, 2, \ldots, m$, $Var(X_i) = \sigma^2, i = 1, 2, \ldots, m$ and $Cov(X_i, X_j) = \rho\sigma^2, i \neq j$. Let the ith observation be $(x_{1i}, \ldots, x_{mi}), i = 1, 2, \ldots, n$. Define

$$\overline{X}_i = \frac{1}{m} \sum_{j=1}^{m} X_{ji},$$

$$W_i = \sum_{j=1}^{m} (X_{ji} - \overline{X}_i)^2,$$

$$B = m \sum_{i=1}^{n} (\overline{X}_i - \overline{X})^2,$$

$$W = W_1 + \cdots + W_n,$$

where B is sum of squares between and W is sum of squares within samples.

   1. Prove (i) $\frac{z_\alpha}{2}\ W \sim (1-\rho)\sigma^2\chi^2(mn-n)$ and (ii) $B \sim (1 + \overline{(m-1)}\rho)\sigma^2\chi^2(n-1)$.

2. Suppose $\frac{(1-\rho)B}{(1+(m-1)\rho)W} \sim F_{(n-1),(mn-n)}$. Prove that when $\rho = 0$, $\frac{W}{W+B}$ follows beta distribution with parameters $\frac{mn-n}{2}$ and $\frac{n-1}{2}$.

22. Show that the joint PDF of $(X_{(1)}, X_{(2)}, \ldots, X_{(n)})$ is given by

$$f(x_{(1)}, \ldots, x_{(n)}) = \begin{cases} n!\Pi_{i=1}^{n} f_{X_i}(x_{(i)}) & \text{if} \quad x_{(1)} \leq x_{(2)} \cdots \leq x_{(n)} \\ 0 & \text{otherwise} \end{cases}.$$

Also, prove that the marginal PDF of $X_{(r)}$ $r = 1, 2, \ldots, n$, where $X_{(r)}$ is the rth minimum, is given by

$$f_{X_{(r)}}(x_{(r)}) = \frac{n!}{(r-1)!(n-r)!}[F(x_{(r)})]^{r-1}[1 - F(x_{(r)})]^{n-r} f(x_{(r)}).$$

23. Let $X_1, X_2, \ldots, X_n$ be a random sample from normal distributed population with mean μ and variance σ^2. Let s^2 be the sample variance. Prove that $\frac{(n-1)s^2}{\sigma^2}$ has χ^2 distribution with $n-1$ degrees of freedom.

24. Let $X_1, X_2, \ldots, X_n$ be a normal distribution with mean $E(X_i) = \mu, i = 1, 2, \ldots, n$, $Var(X_i) = \sigma^2, i = 1, 2, \ldots, n$ and $Cov(X_i, X_j) = \rho\sigma^2, i \neq j$. Let $\overline{X}$ and s^2 denote the sample mean and sample variance, respectively. Then, prove that $\frac{(1-\rho)t}{1+(n-1)\rho}$ has student's t-distribution with $n-1$ degrees of freedom.

25. A sample of 10 claims in an insurance company had mean and variance of 5,478 and 1,723 respectively. On reconciliation, it was found that one claim of 3,250 was wrongly written as 4,250. Calculate the sample mean and standard deviation of the sample with correct values.

26. Government of Tamilnadu wants to analyze the number of children in families for improving their immunization program. They analyze a group of 200 families and report their findings in the form of a frequency distribution shown below:

Number of children	0	1	2	3	4	5	6	7
Number of families	18	30	72	43	25	8	3	1

(a) Calculate sample mean, sample variance and coefficient of variation of the data.

(b) Calculate Coefficient of kurtosis and Coefficient of skewness in the above data.

27. An insurance company wants to analyze the claims for damage due to fire on its household content's policies. Following are the values for a sample of 50 claims in Rupees.

57000	115000	119000	131000	152000	167000	188000	190000	197000	201000
206000	209000	213000	217000	221000	229000	247000	250000	252000	253000
257000	257000	258000	259000	260000	261000	262000	263000	267000	271000
277000	285000	287000	305000	307000	309000	311000	313000	317000	321000
322000	327000	333000	351000	357000	371000	399000	417000	433000	499000

The table given below displays the grouped frequency distribution for the above data.

Claim Size (in 1000's of Rupees)	Frequency
50-99	1
100-149	3
150-199	5
200-249	8
250-299	16
300-349	10
350-399	4
400-449	2
450- 500	1

(a) Calculate the sample geometric mean.

(b) Calculate the sample standard deviation and sample variance.

(c) Calculate the coefficient of variation for the above data.

28. Let $X_1, X_2, \dots, X_5$ be a random sample of size 5 from a normal distribution with mean zero and standard deviation 2. Find the sampling distribution of the statistic $X_1 + 2X_2 - X_3 + X_4 + X_5$.

29. Let X_1, X_2, X_3 be a random sample of size 3 from a standard normal distribution. Find the distribution of $X_1^2 + X_2^2 + X_3^2$. If possible, find the sampling distribution of $X_1^2 - X_2^2$. If not, justify why you can not determine it's distribution.

30. Let $X_1, X_2, \dots, X_6$ be a random sample of size 6 from a standard normal distribution. Find the sampling distribution of the statistics $\frac{X_1+X_2+X_3}{\sqrt{X_4^2+X_5^2+X_6^2}}$ and $\frac{X_1-X_2-X_3}{\sqrt{X_4^2+X_5^2+X_6^2}}$.

31. Let X_1, X_2, X_3 be a random sample of size 3 from an exponential distribution with a parameter $\theta > 0$. Find the distribution of the sample (that is the joint distribution of the r.v. X_1, X_2, X_3).

32. Let $X_1, X_2, \dots, X_n$ be a random sample of size n from a normal population with mean μ and variance $\sigma^2 > 0$. What is the expected value of the sample variance $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$?