Exercises for Lectures 3 and 4

1. Assume there are *m* seeds of maize. We found that *n* of them germinated. Calculate the maximum likelihood estimation of germination percentage *p*.

2. $X_1, X_2, \ldots, X_n$ are random samples from a Normal distribution population $N(\mu, \sigma^2)$. μ and σ^2 are unknown parameters ($-\infty < \mu < +\infty, \sigma^2 > 0$). Calculate the maximum likelihood estimation of μ and σ^2.

3. $X_1, X_2, \ldots, X_n$ are random samples from a Uniform distribution $p(x;\theta) = \begin{cases} \frac{1}{\theta}, 0 < x \le \theta \\ 0, \text{otherwise} \end{cases}$. Use the moment method to calculate the estimator of θ.

4. The mean and variance of the population *X* are μ and σ^2. Is $s_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ an unbiased estimator of σ^2?

5. In 36 cotton fields, the average yield $\bar{x}$ is 4.1 kg. Assume $X \sim N(\mu, \sigma^2)$ and σ^2 is 0.09. Calculate the confidence interval of *X* under the degree of confidence 1-α=99%.

6. From previous experience, we know that the variance of yield per m^2 of a wheat variety is 0.4 (kg^2). Now there are two fields A and B. For 12 samples from A, the mean of yield is $\bar{x}_1 = 1.2 kg/m^2$; for 8 samples from B, the mean of yield is $\bar{x}_2 = 1.4 kg/m^2$. Determine if the yields in A and B are different significantly.

7. There are two groups of mice A and B. Give them different feed and observe the weight increasing *X* of them. X_1 for group A (12 mice): 83, 146, 119, 104, 120, 161, 107, 134, 115, 129, 99, 123; X_2 for group B (7 mice): 70, 118, 101, 85, 107, 132, 94. Determine if the variance of *X* in A and B are different significantly.

8. The nitrogen content in blood plasma of mice is 0.95g/ml. Now we test 9 mice and get their blood plasma: 0.98, 0.84, 0.98, 0.87, 0.89, 0.82, 0.93, 0.91, and 0.88. Is the blood plasma of these mice regular?

9. The variance of heading date of wheat which is cause by environmental factors is

$\sigma_0^2 = 2.781$. Now we test 58 F_2 individuals of wheat and find that the sample variance is $s^2 = 6.458$. Is the variance caused by environment?

10. In a district, the rate of filariasis is 2.7%. Now we surveyed 278 persons in this district and found that 10 of them had filariasis. Is the morbidity regular?

11. $X_1, X_2, \ldots, X_n$ are random samples from a Uniform distribution

$p(x;\theta) = \begin{cases} \frac{1}{\theta}, 0 < x \le \theta \\ 0, \text{otherwise} \end{cases}$. Give the maximum likelihood estimation of θ.

12. μ is the expectation of population X. Prove that the sample mean $\hat{\mu}_1$ and weighted mean $\hat{\mu}_2$ are unbiased.

13. Test a new medicine in 10 patients of silicosis. The hemoglobin concentration before and after treatment is as follows:

Patient	1	2	3	4	5	6	7	8	9	10
Before (x_1)	11.3	15.0	15.0	13.5	12.8	10.0	11.0	12.0	13.0	12.3
After (x_2)	14.0	13.8	14.0	13.5	13.5	12.0	14.7	11.4	13.8	12.0

Can the medicine cause the change of hemoglobin concentration?

14. In two provinces Qinghai and Inner Mongolia, the average of weights from 153 men in Qinghai is 57.41 with standard error 5.77; the average of weights from 686 men in Qinghai is 55.95 with standard error 5.17. Are the weights in these two provinces the same?

15. Two kinds of insecticides A and B. A killed 460 insects in 700; B killed 364 in 500. Are the effects of the two insecticides the same?

16. Use function RAND() in EXCEL to generate 100 sets of 10 pseudo-random numbers of uniform distribution U(0, 1). (1) For each set of 10 random numbers, calculate the sample mean and sample variance. (2) Draw the frequency distribution of the 100 sample means.

Note: For exercises 17-20, you may use the following two theorems.

Theorem 1: Given two independent random numbers X_1 and X_2 from the uniform distribution U(0, 1), transformations $Y_1 = \sqrt{-2\ln(X_1)}\sin(2\pi X_2)$ and $Y_2 = \sqrt{-2\ln(X_1)}\cos(2\pi X_2)$ will generate two independent random numbers Y_1 and Y_2 with the standard normal distribution *N*(0, 1).

Theorem 2: Given one random number *X* from the standard normal distribution *N*(0, 1), transformation $Y = \sigma X + \mu$ will generate one random numbers *Y* with the normal distribution $N(\mu, \sigma^2)$.

17. In EXCEL, generate 100 sets of 5 pseudo-random numbers of the standard normal distribution *N*(0, 1). (1) For each set of 5 random numbers, calculate the sample mean and sample variance. (2) Draw the frequency distribution of the 100 sample means ($\overline{X}$) and 100 sample variances (S^2). (3) Compare the distribution of sample mean with the normal distribution *N*(0, 0.2). (4) Draw the Q-Q plot of the 100 sample means with *N*(0, 0.2).

18. In EXCEL, generate 100 sets of 5 pseudo-random numbers of the standard normal distribution *N*(0, 1). (1) For each set of 5 random numbers, calculate a statistic $\chi^2 = \sum_{i=1,\cdots,5} y_i^2$ which is the sum of square for the 5 random numbers y_1, …, y_5. (2) Draw the frequency distribution of the 100 sums of squares. (3) Compare the distribution of sum of squares with the chi-square distribution $\chi^2(df = 5)$.

19. In Excel, generate 100 sets of 5 pseudo-random numbers of the normal distribution *N*(10, 10). (1) For each set of 5 random numbers, calculate the sample mean and sample variance. (2) For each set of 5 random numbers, calculate the *t* statistic. (3) Draw the frequency distribution of the 100 *t* statistics. (4) Compare the distribution of *t* statistics with the *t* distribution $t(df=4)$.

20. In Excel, generate 100 sets of 6 pseudo-random numbers of the normal distribution *N*(10, 10), and the other 100 sets of 10 pseudo-random numbers of the same distribution. (1) For each set of 6 random numbers and each set of 10 random number, calculate the *F* statistic. (2) Draw the frequency distribution of the 100 *F* statistics. (3) Compare the distribution of *F* statistics with the *F* distribution $F(df_1=5, df_2=9)$.

21. Assume X is a random sample from a Binomial population with *n*=10, and an unknown *p*. Two statistical hypotheses to be tested are H0: *p*=0.5 versus Ha:

$p \neq 0.5$. We reject H0 when X=0, 1, 9, or 10, and we accept H0 when X=2, 3, 4, 5, 6, 7 or 8. (1) Find the type I error α of the statistical test. (2) Given p=0.75, find the Type II error β, and the statistical power of the test. (3) Given p=0.90, find the Type II error β, and the statistical power.

22. Assume X is a random sample from a Binomial population with n=10, and an unknown p. Two statistical hypotheses to be tested H0: p=0.5 versus Ha: $p \neq 0.5$. We reject H0 when X=0, 1, 2, 8, 9, or 10, and we accept H0 when X=3, 4, 5, 6 or 7. (1) Find the type I error α of the statistical test. (2) Given p=0.75, find the Type II error β, and the statistical power of the test. (3) Given p=0.90, find the Type II error β, and the statistical power.