1. For one-way ANOVA test,

   (1) Show that the statistic $(\bar{Y}_{.1},\dots,\bar{Y}_{.k},S_p^2)$ is sufficient for $(\mu_1,\dots,\mu_k,\sigma^2)$

   (2) Show that the F test of one-way ANOVA is a GLRT.

2. When we do two sample test, usually we would have unequal variances. However, in this case, making inference on means is so difficult, which is famous as Behrens-Fisher Problem. The test is described as

$$H_0:\mu_X=\mu_Y \quad \text{versus} \quad H_1:\mu_x\neq\mu_Y$$

   where $X=(X_1,\dots,X_n)$ and $Y=(Y_1,\dots,Y_m)$ are independent. And some statisticians discovered several solutions to the problem.

   (1) Satterthwhite offers an approximation using t distribution. The test statistic is

$$T=\frac{\bar{X}-\bar{Y}}{\sqrt{\frac{S_X^2}{n}+\frac{S_Y^2}{m}}}$$

   Show that under H_0 is true, $T\sim t_\nu$ approximately, where

$$\nu=\frac{(S_X^2/n+S_Y^2/m)^2}{\frac{S_X^4}{n^2(n-1)}+\frac{S_Y^4}{m^2(m-1)}}$$

   (2) Sprott and Farewell notice that if the variance ratio is known, another t statistic can be derived. Suppose $X_i\sim N(\mu_X,\sigma^2)$ $Y_j\sim N(\mu_Y,\rho^2\sigma^2)$ where ρ^2 is known. Show that

$$T=\frac{(\bar{X}-\bar{Y})-(\mu_X-\mu_Y)}{\sqrt{\frac{1}{n}+\frac{\rho^2}{m}}\sqrt{\frac{(n-1)S_X^2+(m-1)S_Y^2/\rho^2}{n+m-2}}}\sim t_{n+m-2}$$

   and furthermore,

$$\frac{S_Y^2}{\rho^2 S_X^2}\sim F_{m-1,n-1}$$

3. Show that for independence testing, when all the parameters are unknown, the observed goodness-of-fit test statistic can be written as

$$d=n(\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{n_{ij}^2}{n_{i.}n_{.j}}-1)$$

   where $n_{i.}=\sum_{j=1}^{c}n_{ij}$ and $n_{.j}=\sum_{i=1}^{r}n_{ij}$

4. $X_1,\dots,X_{n_1},Y_1,\dots,Y_{n_2},Z_1,\dots,Z_{n_3}$ are three independent normal samples with the same unknown variance σ^2. Find a test statistic to test $H_0:\mu_X+\mu_Y=2\mu_Z$ and derive both the rejection region at the significance level α and $(1-\alpha)100\%$ confidence interval for $\mu_X+\mu_Y-2\mu_Z$.

5. A survey shows that the color-blindness may be related to gender. A geneticist constructs a model described as

	Normal	**Color-blindness**
Male	$p/2$	$(1-p)/2$
Female	$p^2/2+p(1-p)$	$(1-p)^2/2$

   If the collected data is

	Normal	Color-blindness
Male	442	38
Female	514	6

do an appropriate goodness-of-fit test at the significance level $\alpha = 0.05$ to test whether the data is consistent with the model.

6. $X_1, \ldots, X_n, Y_1, \ldots, Y_m, Z_1, \ldots, Z_n$ are three independent normal samples with sample mean $\bar{x} = 16, \bar{y} = 14, \bar{z} = 21$ and sample standard deviation $s_X = 12, s_Y = 9, s_Z = 14$.

   (1) If $n = 8, m = 6$, do a two-sided test to test $H_0 : \sigma_X^2 = \sigma_Y^2$ at the significance level $\alpha = 0.05$

   (2) Based on the test result of (1), do a two-sided t test for $H_0 : \mu_X = \mu_Y$ at the significance level $\alpha = 0.05$.

   (3) Suppose the hypothesis in (1) is not rejected, do a one-way ANOVA for $H_0 : \mu_X = \mu_Y$ at the significance level $\alpha = 0.05$ and figure out the relationship of the test statistic in (3) and that in (2).

   (4) If $n = m = 10$, do a one-way ANOVA for $H_0 : \mu_X = \mu_Y = \mu_Z$ at the significance level $\alpha = 0.05$.

   (5) Based on the test result of (4), find the Tukey's interval for all the comparisons where $\alpha = 0.05$.