Chapter 9: Summary

1. Problem: Testing a null hypothesis (H_0) against some alternative (H_A).

2. A simple hypothesis completely specifies a distribution. A composite hypothesis does not completely specify a distribution.

3. A hypothesis can be either one-sided or two-sided.

4. A hypothesis testing problem has essentially four ingredients: (1) null hypothesis, (2) alternative hypothesis, (3) test statistic. and (4) rejection region.

5. Type I error is the probability of rejecting a null hypothesis when it is true. Type II error is the probability of accepting a null hypothesis when it is not true. For a simple alternative, power of a test equals 1 minus the probability of Type II error.

6. For a simple null, the significance level is the probability of Type I error. For a composite null, the significance level is the maximum probability of Type I error under all parameter points corresponding to H_0.

7. Neyman Pearson Lemma: Consider testing a simple null against a simple alternative. Let f_0 and f_A denote respectively the pdf (or pf) under the null and the alternative respectively. Then reject when $f_0(X_1, \cdots, X_n)/f_A(X_1, \cdots, X_n) < c$, where the constant c is determined from a specified significance level α.

8. More generally, if the parameter space $\Omega = (\omega_0) \bigcup (\omega_1)$, and one tests $H_0 : \theta \in \omega_0$ against $H_1 : \theta \in \omega_1$, then a direct generalization of the Neyman Pearson test is given by $\Lambda^* = \max_{\theta \in \omega_0} / \max_{\theta \in \omega_1}$, and reject for small values of Λ^*. However, people prefer to work with $\Lambda = \max_{\theta \in \omega_0} / \max_{\theta \in \Omega}$, and reject for $\Lambda \leq \lambda_0$, where λ_0 is so determined that the test has significance level α. This is known as the generalized likelihood ratio test (GLRT), and Λ is called the GLRT criterion.

9. $\Lambda = \min(\Lambda^*, 1)$. Also, asymptotically, $-2\log_e(\Lambda)$ has a chi-square distribution under H_0.

10. For goodness of fit problems, typically we have several classes with observed frequencies O_i. Under a given hypothesis, we find the corresponding expected frequencies E_i. Then $-2\log_e(\Lambda) = 2\sum_i O_i \log(O_i/E_i)$. Under H_0, $-2\log_e(\Lambda)$ has asymptotically a chisquare distribution with degrees of freedom=no. of classes-no. of parameters estimated under H_0-1. Reject for large values of this statistic at a given significance level.

11. An alternative to the above statistic is the Pearsonioan chisquare given by $\sum_i (O_i - E_i)^2/E_i$. This statistic has the same asymptotic distribution, although the two statistics could be numerically different. Reject for large values of this statistic at a given significance level.

12. For a given observed value of a statistic, the p-value is the probability of observing any value more extreme than the given one. Let us denote the p-value by p^*. For a given significance level α, reject if and only if $p^* < \alpha$.