Suppose the $k$-vector $\mathbf{X}$ is distributed as multinomial $(n, p)$, and we wish to test the null hypothesis $H_0 : p = p^0$ against the alternative $H_1 : p \neq p^0$ using the Pearson chi-square test. We are given a sequence of specific alternatives $p^{(n)}$ satisfying $\sqrt{n}(p^{(n)} - p^0) \to \delta$ for some constant matrix $\delta$. Note that this means $\sum_{i=1}^{k} \delta_i = 0$, a fact that will be used later. Our task is to derive the limit of the power of the test under the sequence of alternatives $p^{(n)}$.

First, we define a noncentral chi-square distribution.

**Definition 25.1** If $A_1, \ldots, A_n$ are independent random variables with $A_i \sim N(\xi_i, 1)$, then the distribution of $A_1^2 + A_2^2 + \cdots + A_n^2$ is noncentral chi-square with $n$ degrees of freedom and noncentrality parameter $\phi = \xi_1^2 + \cdots + \xi_n^2$. We denote this distribution $\chi^2_n(\phi)$. Equivalently, we can say that if $A \sim N_n(\xi, I)$, then $A^T A \sim \chi^2_n(\phi)$ where $\phi = \xi^T \xi$.

Note that this is not a valid definition unless we may prove that the distribution of $A^T A$ depends on $\xi$ only through $\phi = \xi^T \xi$. We prove this as follows. First, note that if $\phi = 0$ then there is nothing to prove. Otherwise, define $\xi^* = \xi / \sqrt{\phi}$. Next, find an orthogonal matrix $Q$ whose first row is $(\xi^*)'$. This may be accomplished by, for example, Gram-Schmidt orthogonalization. Then $QA \sim N_k(Q\xi, I)$. Since $Q\xi$ is a vector with first element $\sqrt{\phi}$ and remaining elements 0, clearly $QA$ has a distribution that depends on $\xi$ only through $\phi$. But $A^T A = (QA)'(QA)$, so this proves that Definition 25.1 is valid.

We will derive the power of the chi-square test by adapting the projection matrix technique of Topic 22. First, we prove a lemma.

**Lemma 25.1** Suppose $Z \sim N_k(\mu, P)$, where $P$ is a projection matrix of rank $r \leq k$ and $P\mu = \mu$.

Then $Z^T Z \sim \chi^2_r(\mu^T \mu)$.

**Proof:** Since $P$ is a covariance matrix, it is symmetric, which means that there exists an orthogonal matrix $Q$ with $QPQ^{-1} = \text{diag}(\Lambda)$, where $\Lambda$ is the vector of eigenvalues of $P$. Since $P$ is a projection matrix, all of its eigenvalues are 0 or 1. Since $P$ has rank $r$, exactly $r$ of the eigenvalues are 1. Without loss of generality, assume that the first $r$ entries of $\Lambda$ are 1 and the last $k - r$ are 0. The random vector $QZ$ is $N_n(Q\mu, \text{diag}(\Lambda))$, which implies that $Z^T Z = (QZ)'(QZ)$ is by definition distributed as $\chi^2_r(\phi) + \varphi$, where

$$\phi = \sum_{i=1}^{r} (Q\mu)_i^2 \quad \text{and} \quad \varphi = \sum_{i=r+1}^{k} (Q\mu)_i^2.$$

Note, however, that

$$Q\mu = QP\mu = QPQ\mu = \text{diag}(\Lambda)Q\mu. \quad (76)$$

Since entries $r + 1$ through $k$ of $\Lambda$ are zero, the corresponding entries of $Q\mu$ must be zero because of equation (76). This implies two things: First, $\varphi = 0$; and second,

$$\phi = \sum_{i=1}^{r} (Q\mu)_i^2 = \sum_{i=1}^{k} (Q\mu)_i^2 = (Q\mu)'(Q\mu) = \mu^T \mu.$$

Thus, $Z^T Z \sim \chi^2_r(\mu^T \mu)$, which proves the result. ■

Define $\Gamma = \text{diag}(p^0)$. Let $\Sigma = \Gamma - p^0(p^0)'$ be the usual multinomial covariance matrix under the null hypothesis; i.e., $\sqrt{n}(X^{(n)} / n - p^0)' \Gamma^{-1} \sim N_k(0, \Sigma)$ if $X^{(n)} \sim \text{multinomial}(n, p^0)$. Consider $X^{(n)}$ to have instead
a multinomial \((n, p^{(n)})\) distribution. Under the assumption made earlier that \(\sqrt{n}(p^{(n)} - p^0) \rightarrow \delta\), it may be shown that
\[
\sqrt{n}(X^{(n)}/n - p^{(n)}) \xrightarrow{L} N_k(0, \Sigma).
\] (77)
This is shown, for example, in Theorem 5.5.3 on p. 327; we omit the details here. We claim that the limit (77) implies that the chi-square statistic
\[n(X^{(n)}/n - p^{(n)}) + \sqrt{n}(p^{(n)} - p^0)\]
shown that
\[n(X^{(n)}/n - p^{(n)})\]
implies that the chi square statistic
\[\chi^2_{k-1}(\delta\Gamma^{-1}\delta)\]
converges to
\[\chi^2_{k-1}(\delta\Gamma^{-1}\delta)\]
Therefore, Slutsky’s theorem implies that \(V^{(n)} \xrightarrow{L} N_k(\delta, \Sigma)\). A further application of Slutsky’s theorem implies
\[
\Gamma^{-1/2}V^{(n)} \xrightarrow{L} N_k(\Gamma^{-1/2}\delta, \Gamma^{-1/2}\Sigma\Gamma^{-1/2}).
\]
Thus, the result we wish to prove follows from Lemma 25.1 if we can demonstrate that \((\Gamma^{-1/2}\Sigma\Gamma^{-1/2})^{-1}(\Gamma^{-1/2}\delta) = (\Gamma^{-1/2}\delta)\). To check this last fact, note that
\[
\Gamma^{-1/2}\Sigma\Gamma^{-1/2}\delta = \Gamma^{-1/2}[\Gamma - \rho^0(\rho^0)^t]\Gamma^{-1}\delta = \Gamma^{-1}\delta - \rho^0(1)^t\delta = \Gamma^{-1/2}\delta
\]
since \(1^t\delta = \sum_{i=1}^k \delta_i = 0\). Thus, we conclude that the chi-square statistic converges in distribution to \(\chi^2_{k-1}(\delta\Gamma^{-1}\delta)\) under the sequence of alternatives \(\rho^{(1)}, \rho^{(2)}, \ldots\).

**Example 25.1** For a particular trinomial experiment with \(n = 200\), suppose the null hypothesis is \(H_0 : p = p^0 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})\). (This hypothesis might arise in the context of a genetics experiment.) We may calculate the approximate power of the Pearson chi-square test at level \(\alpha = 0.01\) against the alternative \(p = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})\).

First, set \(\delta = \sqrt{n}(p - p^0) = \sqrt{200}(\frac{1}{3}, -\frac{1}{3}, \frac{1}{3})\). Under the alternative \(p\), the chi square statistic is approximately noncentral \(\chi^2_{2}\) with noncentrality parameter
\[
\frac{\delta^t \text{diag}(p^0)^{-1}\delta}{2} = 200 \left(\frac{4}{144} + \frac{2}{36} + \frac{4}{144}\right) = \frac{200}{9}.
\]
Since the test rejects \(H_0\) whenever the statistic is larger than the .99 quantile of \(\chi^2_{2}\), namely 9.210, the power is approximated by \(P(\chi^2_{2}(\frac{200}{9}) > 9.210) \approx 0.965\). These values were found using R as follows:
\[
> \text{qchisq}(0.99, 2)
\]
[1] 9.21034
\[
> \text{1-pchisq}(\text{Last.value}, 2, \text{ncp}=200/9)
\]
[1] 0.965006

**Problems**

**Problem 25.1** *Hotelling’s T^2*. Suppose \(X^{(1)}, X^{(2)}, \ldots\) are iid from some \(k\)-dimensional distribution with mean \(\mu\) and finite nonsingular covariance matrix \(\Sigma\). Let \(S_n\) denote the sample covariance matrix
\[
S_n = \frac{1}{n-1} \sum_{j=1}^n (X^{(j)} - \overline{X})(X^{(j)} - \overline{X})^t.
\]
To test $H_0 : \mu = \mu^0$ against $H_1 : \mu \neq \mu^0$, define the statistic

$$T^2 = (\mathbf{V}(n))^t S_n^{-1} (\mathbf{V}(n)),$$

where $\mathbf{V}(n) = \sqrt{n}(\mathbf{X} - \mu^0)$. This is called Hotelling’s $T^2$ statistic.

[Notes: This is a generalization of the square of a unidimensional t statistic. If the sample is multivariate normal, then $((n - k)/(nk - k))T^2$ is distributed as $F_{k,n-k}$. A Pearson chi square statistic may be shown to be a special case of Hotelling’s $T^2$.]

(a) You may assume that $S_n^{-1} \overset{P}{\to} \Sigma^{-1}$ (this follows from the WLLN since $P(S_n$ is nonsingular) $\to 1$). Prove that under the null hypothesis, $T^2 \overset{L}{\to} \chi^2_k$.

(b) Let $\{\mu^{(n)}\}$ be alternatives such that $\sqrt{n}(\mu^{(n)} - \mu^0) \overset{L}{\to} \delta$. You may assume that under $\{\mu^{(n)}\}$,

$$\sqrt{n}(\mathbf{X} - \mu^{(n)}) \overset{L}{\to} N_k(0, \Sigma).$$

Find (with proof) the limit of the power against the alternatives $\{\mu^{(n)}\}$ of the test that rejects $H_0$ when $T^2 \geq c_\alpha$, where $P(\chi^2_k > c_\alpha) = \alpha$.

(c) An approximate $1 - \alpha$ confidence set based on the result in part (a) may be formed by plotting the elliptical set

$$\{\mu : n(\mathbf{X} - \mu)^t S_n^{-1}(\mathbf{X} - \mu) = c_\alpha\}.$$

For a random sample of size 100 from $N_2(0, \Sigma)$, where $\Sigma = \begin{pmatrix} 1 & 3/5 \\ 3/5 & 1 \end{pmatrix}$, produce a scatterplot of the sample and plot 90% and 99% confidence sets on this scatterplot.

Hints: In part (c), to produce a random vector with the $N_2(0, \Sigma)$ distribution, take a $N_2(0, I)$ random vector and left-multiply by a matrix $A$ such that $AA^t = \Sigma$. It is not hard to find such an $A$ (it may be taken to be lower triangular). One way to graph the ellipse is to find a matrix $B$ such that $B^t S_n^{-1} B = I$. Then note that

$$\{\mu : n(\mathbf{X} - \mu)^t S_n^{-1}(\mathbf{X} - \mu) = c_\alpha\} = \{\mathbf{X} - B\mathbf{\nu} : \mathbf{\nu}^t \mathbf{\nu} = c_\alpha/n\},$$

and of course it’s easy to find points $\mathbf{\nu}$ such that $\mathbf{\nu}^t \mathbf{\nu}$ equals a constant. To find a matrix $B$ such as the one specified, note that the matrix of eigenvalues of $S_n$, properly normalized, gives an orthogonal matrix that diagonalizes.

Problem 25.2 Suppose we have a tetranomial experiment and wish to test $H_0 : \mathbf{p} = (1/4, 1/4, 1/4, 1/4)$ against $H_1 : \mathbf{p} \neq (1/4, 1/4, 1/4, 1/4)$ at the .05 level.

(a) Approximate the power of the test against the alternative $(1/10, 2/10, 3/10, 4/10)$ for a sample of size $n = 200$.

(b) Give the approximate sample size necessary to give power of 80% against the alternative in part (a).

Hints: You can use the Splus function `pchisq` directly in part (a), but in part (b) you may have to use a trial-and-error approach with `pchisq`. 
26. The Wilcoxon Rank-Sum Test

Lehmann §3.4

Suppose that $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ are two independent iid samples, with

$$P(X_i \leq t) = P(Y_j \leq t + \theta) = F(t)$$

for some continuous cdf $F(t)$ with $F'(t) = f(t)$. Thus, the distribution of the $Y_j$ is shifted by $\theta$ from the distribution of the $X_i$. We wish to test $H_0: \theta = 0$ against $H_1: \theta > 0$, so clearly $\theta_0 = 0$ in what follows.

To do the asymptotics here, we will assume that $n$ and $m$ are actually both elements of separate sequences of sample sizes, indexed by a third variable, say $k$. Thus, $m = m_k$ and $n = n_k$ both go to $\infty$ as $k \to \infty$, and we suppress the subscript $k$ on $m$ and $n$ for convenience of notation. Suppose that we combine the $X_i$ and $Y_j$ into a single sample of size $m + n$. Define the Wilcoxon rank-sum statistic to be

$$W_k = \sum_{j=1}^n \text{Rank of } Y_j \text{ among combined sample}.$$

Letting $Y_{(1)}, \ldots, Y_{(n)}$ denote the order statistics for the sample of $Y_j$ as usual, we may rewrite $W_k$ in the following way:

$$W_k = \sum_{j=1}^n \text{Rank of } Y_{(j)} \text{ among combined sample}$$
$$= \sum_{j=1}^n \left( j + \#\{i: X_i < Y_{(j)}\} \right)$$
$$= \frac{n(n+1)}{2} + \sum_{j=1}^n \sum_{i=1}^m I\{X_i < Y_{(j)}\}$$
$$= \frac{n(n+1)}{2} + \sum_{j=1}^n \sum_{i=1}^m I\{X_i < Y_j\}. \quad (79)$$

Let $N = N_k$ denote the combined sample size $m + n$, and suppose that $m/N \to \rho$ as $k \to \infty$ for some constant $\rho \in (0, 1)$. For a sequence of alternatives $\theta_1, \theta_2, \ldots$, suppose that $\sqrt{N}(\theta_k - \theta_0) \to \Delta$ for a positive, finite constant $\Delta$. Setting

$$\mu(\theta) = E_\theta W_k \quad \text{and} \quad \tau(\theta) = \sqrt{N \text{Var } W_k},$$

we obtain

$$\mu(\theta_0) = \frac{n(n+1)}{2} + \frac{mn}{2} = \frac{n(N+1)}{2}. \quad (75)$$

To evaluate $\tau(\theta_0)$, let $Z_j = \sum_{i=1}^m I\{X_i < Y_j\}$. Then the $Z_j$ are identically distributed but not independent, and we have $E_0 Z_j = n/2$.

$$\text{Var}_0 Z_j = \frac{n}{4} + \frac{n(n-1)}{3} - \frac{n(n-1)}{4}$$
$$= \frac{n(n+2)}{12}.$$
and
\[ E_{\theta_0} Z_i Z_j = \sum_{r=1}^{n} \sum_{s=1}^{n} P_{\theta_0} (X_r < Y_i \text{ and } X_s < Y_j) = \frac{n(n-1)}{4} + n P_{\theta_0} (X_1 < Y_i \text{ and } X_1 < Y_j). \]

Therefore, we obtain
\[ \text{Cov}_{\theta_0} (Z_i, Z_j) = \frac{n(n-1)}{4} + \frac{n^2}{4} = \frac{n}{12}, \]

so
\[ \tau^2(\theta_0) = Nm \text{ Var } Z_1 + Nm(m-1) \text{ Cov } (Z_1, Z_2) = \frac{Nmn(n+2)}{12} + \frac{Nm(m-1)n}{12} = \frac{Nmn(N+1)}{12}. \]

It is possible to show that
\[ \frac{\sqrt{N} \{W_k - \mu(\theta_0)\}}{\tau(\theta_0)} \leq N(0,1) \] under \( H_0 \) and
\[ \frac{\sqrt{N} \{W_k - \mu(\theta_k)\}}{\tau(\theta_0)} \leq N(0,1) \] under the alternatives \( \{\theta_k\} \). However, we do not do so here. See Section 2.8 of Lehmann for details on (80) and Section 3.4 for details on (81).

To find the limiting power of the rank-sum test, we may use Theorem 23.1 to conclude that
\[ \beta_k(\theta_k) \rightarrow \lim_{k \to \infty} \Phi \left( \frac{\Delta \mu(\theta_0)}{\tau(\theta_0)} - u_\alpha \right). \] (82)

Thus, we should evaluate \( \mu'(\theta) \). To this end, note that
\[ P_{\theta} (X_1 < Y_1) = E_{\theta} \{P_{\theta} (X_1 < Y_1 \mid Y_1)\} = E_{\theta} F(Y_1) = \int_{-\infty}^{\infty} F(y) f(y - \theta) \, dy = \int_{-\infty}^{\infty} F(y + \theta) f(y) \, dy. \]

Therefore,
\[ \frac{d}{d\theta} P_{\theta} (X_1 < Y_1) = \int_{-\infty}^{\infty} f(y + \theta) f(y) \, dy. \]

This gives
\[ \mu'(0) = mn \int_{-\infty}^{\infty} f^2(y) \, dy. \]

Thus, the efficacy of the Wilcoxon rank-sum test is
\[ \lim_{k \to \infty} \frac{\mu'(\theta_0)}{\tau(\theta_0)} = \lim_{k \to \infty} \frac{mn \sqrt{12} \int_{-\infty}^{\infty} f^2(y) \, dy}{\sqrt{mnN(N+1)}} = \sqrt{\frac{12 \rho(1-\rho)}{12 \rho(1-\rho)^2}} \int_{-\infty}^{\infty} f^2(y) \, dy. \]

The asymptotic power of the test follow immediately from (82).
Problems

Problem 26.1 In the situation of equation (78), suppose Var \( X_i = \sigma^2 < \infty \) and we wish to test the hypotheses \( H_0 : \theta = 0 \) vs. \( H_1 : \theta > 0 \) using the two-sample Z-statistic

\[
\frac{\bar{X} - \bar{Y}}{\sigma \sqrt{\frac{1}{m} + \frac{1}{n}}}
\]

Note that this Z-statistic is \( s/\sigma \) times the usual T-statistic, so the asymptotic properties of the T-statistic are the same as those of the Z-statistic.

(a) Find the efficacy of the Z test. Justify your use of Theorem 23.1.

(b) Find the ARE of the Z test with respect to the rank-sum test for normally distributed data.

(c) Find the ARE of the Z test with respect to the rank-sum test if the data come from a double exponential distribution with \( f(t) = \frac{1}{2\lambda}e^{-|t/\lambda|} \).

(d) Prove that the ARE of the Z-test with respect to the rank-sum test can be arbitrarily close to zero.

Hint: In part (d), it suffices to take \( \epsilon > 0 \) and find an example for which the ARE is less than \( \epsilon \).