1. (a) State the Borel-Cantelli Lemma and its converse.

(b) Let \(X_1, X_2, \ldots\) be i.i.d. from a distribution with density,
\[ f(x) = \theta x^{-(\theta+1)} \]
on the interval \((1, \infty)\). For what value of \(\theta\) is it true that \((1/n)X_n \xrightarrow{a.s.} 0\).

2. Let \(X_1, X_2, \ldots\) be independent random variables with \(X_k\) having the distribution
\[ X_k = \begin{cases} \frac{1}{\sqrt{k}} & \text{with probability } \frac{\sqrt{k}}{(k+1)^{1/2}} \\ -1 & \text{with probability } \frac{1}{(k+1)^{1/2}} \end{cases} \]

(a) Let \(S_n = \sum_{k=1}^{n} X_k\). Find \(E(S_n)\) and \(\text{Var}(S_n)\). Note that \(\text{Var}(S_n) \sim 2\sqrt{n}\).

(b) Check that the UAN condition holds.

(c) Show whether or not \((S_n - E(S_n))/\sqrt{\text{Var}(S_n)}\) converges in law to the standard normal distribution by checking the Lindeberg condition.

3. Suppose we are given \(n\) independent trials resulting in \(c\) possible cells, each trial having probability \(p_i\) of falling in cell \(i\), for \(i = 1, \ldots, c\). Let \(n_i\) denote the number of trials falling in cell \(i\).

(a) What is Pearson’s chi-square for testing the hypothesis that the true probabilities are \(p_i\) for \(i = 1, \ldots, c\)?

(b) Find the transformed chi-square with the transformation, \(g(p) = \log(p)\) applied to each cell. Find the modified transformed chi-square.

(c) What is the approximate large sample distribution of the modified transformed chi-square if the true cell probabilities are \(p_i^0\) for \(i = 1, \ldots, c\)?

4. In sampling from a population of \(N\) objects having values \(z_1, z_2, \ldots, z_N\), first a sample of size \(n < N/2\) is taken without replacement. Later a second sample of size \(n\) is taken from the remaining \(N - n\) objects without replacement. The difference of the means of the two samples is used to compare the samples. This leads to a rank statistic of the form \(S_N = \sum_{j=1}^{N} z_j a(R_j)\), where \(a(i) = 1\) for \(i = 1, \ldots, n\), \(a(i) = -1\) for \(i = n + 1, \ldots, 2n\), and \(a(i) = 0\) for \(i = 2n + 1, \ldots, N\).

(a) What are the mean and the variance of \(S_N\)?

(b) Assume that \(n \to \infty\) as \(N \to \infty\). Under what condition on the \(z_i\) is it true that \((S_N - E(S_N))/\sqrt{\text{Var}(S_N)} \xrightarrow{L} \mathcal{N}(0,1)\)?
5. (a) Give the definition of the Kullback-Leibler Information number, \( K(f_0, f_1) \).
(b) What is the Information Inequality?
(c) Suppose \( f_0(x) \) is the density of the binomial distribution, \( \mathcal{B}(n, 1/2) \) (with sample size \( n \) and probability of success \( 1/2 \)), and \( f_1(x) \) is the density of the binomial distribution, \( \mathcal{B}(n, 3/4) \). Find \( K(f_0, f_1) \) and check that the inequality holds.

6. Let \((X_1, Y_1), \ldots, (X_n, Y_n)\) be a sample from a bivariate distribution with density
\[
f(x, y|\mu, \theta) = \theta^2 \mu x \exp\{-\theta x(1 + \mu y)\} \quad \text{for } x > 0 \text{ and } y > 0,
\]
where \( \mu > 0 \) and \( \theta > 0 \) are parameters.
(a) Find the maximum likelihood estimates of \( \mu \) and \( \theta \).
(b) Find the Fisher Information matrix for this distribution.
(c) What is the asymptotic distribution of the MLE of \( \mu \) when \( \theta \) is unknown? What is the asymptotic distribution of the MLE of \( \mu \) when \( \theta \) is known?

7. Let \( X_1, \ldots, X_n \) be a sample from the Poisson distribution \( \mathcal{P}(\lambda) \), let \( Y_1, \ldots, Y_n \) be a sample from a Poisson distribution, \( \mathcal{P}(\lambda + \beta_1) \), let \( Z_1, \ldots, Z_n \) be a sample from the Poisson distribution, \( \mathcal{P}(\lambda + \beta_2) \), with all three parameters, \( \lambda, \beta_1, \beta_2 \), unknown. Assume that all three samples are independent.
(a) Find the likelihood ratio test statistic for testing the hypothesis \( H_0: \beta_1 = \beta_2 \).
(b) What function of the likelihood ratio test statistic has asymptotically a chi-square distribution, and how many degrees of freedom does it have in this case?

8. A sample of size \( n \) is taken in a multinomial experiment with \( c^2 \) cells denoted \((i, j)\), \( i = 1, \ldots, c \) and \( j = 1, \ldots, c \). Let \( p_{ij} \) denote the probability of cell \((i, j)\), and let \( n_{ij} \) denote the number falling in cell \((i, j)\), so that \( \sum \sum p_{ij} = 1 \) and \( \sum \sum n_{ij} = n \).
(a) Let \( H \) denote the hypothesis of symmetry, that \( p_{ij} = p_{ji} \) for all \( i \) and \( j \). Find the chi-square test of \( H \) against all alternatives? How many degrees of freedom does it have?
(b) Let \( H_0 \) denote the hypothesis that all off-diagonal elements are equal: \( p_{ij} = q \) for all \( i \neq j \), for some \( q \). Note that under \( H_0 \), \( p_{11} + p_{22} + \ldots + p_{cc} + c(c-1)q = 1 \). Find the chi-square test of \( H_0 \) against all alternatives. How many degrees of freedom?
(c) What, then, is the chi-square test of \( H_0 \) against \( H \), and how many degrees of freedom does it have?
1. (a) If $A_1, A_2, \ldots$ are events such that $\sum_{j=1}^{\infty} P(A_j) < \infty$, then $P(A_n \ i.o.) = 0$. Conversely, if the $A_j$ are independent events, and $\sum_{j=1}^{\infty} P(A_j) = \infty$, then $P(A_n \ i.o.) = 1$.

(b) Let $\epsilon$ be an arbitrary positive number. Then $(1/n)X_n \overset{a.s.}{\to} 0$ if, and only if, $P((1/n)X_n > \epsilon \ i.o.) = 0$. Since
\[ \sum_{n=1}^{\infty} P((1/n)X_n > \epsilon) = \sum_{n=1}^{\infty} P(X_n > n\epsilon) = \sum_{n=1}^{\infty} 1/(n\epsilon)^\theta < \infty \]
if, and only if, $\theta > 1$, we have $(1/n)X_n \overset{a.s.}{\to} 0$ if, and only if, $\theta > 1$.

2. (a) $E(X_k) = 0$ and $\text{Var}(X_k) = 1/\sqrt{k}$. So $E(S_n) = 0$ and $B_n^2 = \text{Var}(S_n) = \sum_{i=1}^{n} 1/\sqrt{k} \sim \int_{1}^{n} (1/x) dx \sim 2\sqrt{n}.$
(b) $\max_{1 \leq j \leq n} 1/\sqrt{j} = 1$, so $\max_j \text{Var}(X_j)/B_n^2 \sim 1/2\sqrt{n} \to 0$.
(c) Since $|X_j| \leq 1$ for all $j$,
\[ \frac{1}{B_n^2} \sum_{j=1}^{n} E(X_j^2 I(X_j^2 > \epsilon^2 B_n^2)) \leq \frac{1}{B_n^2} \sum_{j=1}^{n} E(X_j^2 I(1 > \epsilon^2 B_n^2)) = \text{I}(1 > \epsilon^2/2\sqrt{n}) = 0 \]
for $n$ sufficiently large. So, $S_n/B_n \overset{L}{\to} N(0,1)$, or $S_n/n^{1/4} \overset{L}{\to} N(0,1/2)$

3. (a) $\chi_p^2 = n \sum_{j=1}^{c} \left(\frac{(n_j/n) - p_j}{p_j}\right)^2$.
(b) $\chi^2 = n \sum_{j=1}^{c} p_j (\log(n_j/n) - \log(p_j))^2$ and $\chi^2_{TM} = \sum_{j=1}^{c} n_j (\log(n_j/n) - \log(p_j))^2$.
(c) The limiting distribution of noncentral $\chi^2_{\lambda-1}(\lambda)$, with $c-1$ degrees of freedom and noncentrality parameter $\lambda = n \sum_{j=1}^{c} p_j^0 (\log(p_j^0) - \log(p_j))^2$.

4. (a) Since $\bar{a}_N = 0$, we have $E S_N = 0$. The variance of $S_N$ is $(N^2/(N-1))\sigma^2 z^2 \sigma^2$, and since $\sigma^2 = (1/N) \sum_{i=1}^{N} a(i)^2 = 2n/N$, we have $\text{Var}(S_N) = (2nN/(N-1))\sigma^2 z^2$.
(b) For asymptotic normality of $S_N$, we need
\[ \max_j (z_j - \bar{z}_N)^2 \max(a(j) - \bar{a}_N)^2 \to 0. \]
We have $\max_j (a(j) - \bar{a}_N)^2 = 1$, and $\sigma^2 = 2n/N$. Then the above condition becomes
\[ \frac{\max_j (z_j - \bar{z}_N)^2}{2n\sigma^2} \to 0. \]

5. (a) $K(f_0, f_1) = E_0 \log \left( \frac{f_0(X)}{f_1(X)} \right)$, where $E_0$ represents the expectation when $f_0(x)$ is the density of $X$. 

Solutions to the Final Examination, Stat 200C, Spring 2010.
(b) $K(f_0, f_1) \geq 0$, with equality if, and only if, $f_0(x)$ and $f_1(x)$ are the same distribution.

(c) $f_0(x) = \binom{n}{x} (1/2)^n$ and $f_1(x) = \binom{n}{x} (3/4)^x (1/4)^{n-x}$, so $f_0(x)/f_1(x) = 2^n/3^x$. So $K(f_0, f_1) = E_0(n \log 2 - X \log 3) = n \log 2 - (n/2) \log 3 = (n/2)[\log 4 - \log 3]$. This is obviously positive.

6. (a) The log-likelihood function is $\ell_n(\theta, \mu) = 2n \log \theta + n \log \mu + \sum_1^n \log x_i - \theta \sum_1^n x_i(1 + \mu y_i)$. 
\[ \frac{\partial \ell_n}{\partial \theta} = (2n/\theta) - \sum_1^n x_i(1 + \mu y_i) = 0 \quad \text{implies} \quad 2 = \theta x + \theta \mu xy \quad \text{and} \quad \frac{\partial \ell_n}{\partial \mu} = (n/\mu) - \theta \sum_1^n x_i y_i = 0 \quad \text{implies} \quad 1 = \theta \mu xy. \]

Solving these equations gives $\hat{\theta} = 1/X_n$ and $\hat{\mu} = X_n/XY_n$, where $XY_n = (1/n) \sum_1^n X_i Y_i$.

(b) $\Psi(x, \theta, \mu) = ((2/\theta) - x(1 + \mu y), (1/\mu) - \theta xy)$, which shows that $E(XY) = 1/\mu \theta$, so that

\[ \hat{\psi} = \begin{pmatrix} -2/\theta^2 & -xy \\ -xy & -1/\mu^2 \end{pmatrix} \quad \text{and} \quad I(\theta, \mu) = -E\hat{\psi} = \begin{pmatrix} 2/\theta^2 & 1/\mu \theta \\ 1/\mu \theta & 1/\mu^2 \end{pmatrix} \]

(c) Since $\text{Det}(I) = 1/\mu^2 \theta^2$, we have $E(XY) = 1/\mu \theta$, so that

\[ I(\theta, \mu)^{-1} = \mu^2 \theta^2 \begin{pmatrix} 1/\mu^2 & -1/\mu \theta \\ -1/\mu \theta & 2/\theta^2 \end{pmatrix} = \begin{pmatrix} \theta^2 & -\mu \theta \\ -\mu \theta & 2\mu^2 \end{pmatrix}. \]

So when $\theta$ is unknown, $\sqrt{n}(\hat{\mu} - \mu) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 2\mu^2)$. When $\theta$ is known, the asymptotic variance of the MLE is the reciprocal of the lower right corner of the information matrix, namely $\mu^2$. So $\sqrt{n}(\hat{\mu} - \mu) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \mu^2)$.

(Here, the MLE of $\mu$ is $\bar{\mu} = 1/(\theta \bar{X}_n)$.)

7. (a) The log-likelihood function is $\ell_n = \log L_n(\lambda, \beta_1, \beta_2) = -(3\lambda + \beta_1 + \beta_2)n + \log(\lambda) \sum X_i + \log(\lambda + \beta_1) \sum Y_i + \log(\lambda + \beta_2) \sum Z_i$ plus a term not involving the parameters. The likelihood equations are

\[ \frac{\partial \ell_n}{\partial \lambda} = -3n + \lambda^{-1} \sum X_i + (\lambda + \beta_1)^{-1} \sum Y_i + (\lambda + \beta_2)^{-1} \sum Z_i \]
\[ \frac{\partial \ell_n}{\partial \beta_1} = -n + (\lambda + \beta_1)^{-1} \sum Y_i \]
\[ \frac{\partial \ell_n}{\partial \beta_2} = -n + (\lambda + \beta_2)^{-1} \sum Z_i \]

The maximum likelihood estimates are $\hat{\lambda} = \bar{X}_n$, $\hat{\beta}_1 = \bar{Y}_n - \bar{X}_n$, and $\hat{\beta}_2 = \bar{Z}_n - \bar{X}_n$. In a similar way, the MLE’s under $H_0$ are $\hat{\lambda} = \bar{X}_n$, and $\hat{\beta}_1 = \hat{\beta}_2 = (\bar{Y}_n + \bar{Z}_n)/2 - \bar{X}_n$. The likelihood ratio test rejects $H_0$ for small values of

\[ \Lambda = \frac{L_n(\hat{\lambda}, \hat{\beta}_1, \hat{\beta}_2)}{L_n(\hat{\lambda}, \hat{\beta}_1, \hat{\beta}_2)} = \frac{(\bar{Y}_n + \bar{Z}_n)/2)^n(\bar{Y}_n + \bar{Z}_n)}{\bar{Y}_n^n \bar{Z}_n^n}. \]

(b) $-2 \log \Lambda$ has asymptotically a chi-square distribution with 1 degree of freedom.
8. (a) Under $H$, the maximum likelihood estimates of the $p_{ij}$ are $\hat{p}_{ii} = n_{ii}/n$ and for $i \neq j$, $\hat{p}_{ij} = (n_{ij} + n_{ji})/2n$. There are $(c - 1) + (c - 2) + \cdots + 1 = c(c - 1)/2$ restrictions going from the general hypothesis to $H$. So the chi-square test of $H$ rejects $H$ if $\chi^2(\hat{p})$ is greater than the appropriate cutoff point for a chi-square distribution with $c(c - 1)/2$ degrees of freedom.

   (b) Under $H_0$, the likelihood is proportional to $[\prod_{j=1}^{n} p_{jj}^{n_{jj}}]q^m$, where $m = n - \sum_{j=1}^{c} n_{jj}$. So the maximum likelihood estimates are
\[
\hat{p}_{jj} = \frac{n_{jj}}{n} \quad \text{and} \quad \hat{q} = [1 - \sum_{1}^{c} \hat{p}_{jj}]/(c(c - 1)).
\]

There are $c$ parameters estimated so the chi-square test of $H_0$ rejects $H_0$ if $\chi^2(\hat{p})$ is greater than the appropriate cutoff point for a chi-square distribution with $(c^2 - 1) - c = c^2 - c - 1$ degrees of freedom.

   (c) The chi-square test of $H_0$ within $H$, rejects $H_0$ if $\chi^2(\hat{p}) - \chi^2(\hat{\hat{p}})$ is greater than the appropriate cutoff point for a chi-square distribution with $c^2 - c - 1 - (c(c - 1)/2) = (c(c - 1)/2) - 1$ degrees of freedom.