1. (a) Define: \(X_n\) converges in probability to \(X\).
(b) Define: \(X_m\) converges in quadratic mean to \(X\).
(c) Show that if \(X_n\) converges in quadratic mean to \(X\), then \(X_n\) converges in probability to \(X\).

2. Let \(X_1, X_2, \ldots\) be independent random variables with \(X_k\) having a uniform distribution over the interval \((0, k)\). Let \(S_n = \sum_{k=1}^{n} X_k\).
   (a) Find \(E(S_n)\) and \(\text{Var}(S_n)\).
   (b) Show that \((S_n - E(S_n))/\sqrt{\text{Var}(S_n)}\) converges in law to the standard normal distribution by checking the Lindeberg condition.

3. Let \(X_1, \ldots, X_n\) be a sequence of i.i.d. random variables taking values in the set \{red, white, blue\} with \(P(X_j = \text{red}) = P(X_j = \text{white}) = P(X_j = \text{blue}) = 1/3\). Let \(S_n\) denote the total number of times that the sequence (red, white, blue) appears in that order from left to right in the sequence. That is, if \(Z_j = \mathbb{I}(X_j = \text{red}, X_{j+1} = \text{white}, X_{j+2} = \text{blue})\), then \(S_n = \sum_{j=1}^{n-2} Z_j\).
   (a) Find \(E(S_n)\) and \(\text{Var}(S_n)\).
   (b) What is the limiting distribution of \((S_n - ES_N)/\sqrt{n}\)?

4. Let \(X_1, \ldots, X_n\) be a sample from the Pareto distribution with distribution function \(F(x) = (1 - (1/x))I(X > 1)\) and density \(f(x) = (1/x^2)I(x > 1)\).
   (a) Find the first quartile and the median of the distribution.
   (b) Find the joint asymptotic distribution of the sample first quartile and the sample median.
   (c) Let \(M_n\) denote the sample maximum. Find the limiting distribution function of \(M_n/b_n\) for suitable choice of \(b_n\) to get a non-degenerate limit.

5. Let \(X_1, \ldots, X_n\) be a sample from the logistic distribution with density
   \[
   f(x|\theta) = \frac{e^{-(x-\theta)}}{(1 + e^{-(x-\theta)})^2}
   \]
   (a) Find the likelihood equation for solving for the maximum likelihood estimate.
   (b) This density is symmetric about \(\theta\), so \(\theta\) is the median. What is the asymptotic distribution of the sample median as an estimate of \(\theta\)?
   (c) Show how to improve the median to a fully efficient estimate using the likelihood equation. (Note: Fisher information is \(\mathcal{I}(\theta) = 1/3\).)
6. (a) Let \((X_1, Y_1), \ldots, (X_n, Y_n)\) be a sample from a bivariate distribution with mean \((\mu_x, \mu_y)\), and covariance matrix, 
\[
\Sigma = \begin{pmatrix}
\sigma_x^2 & \sigma_{xy} \\
\sigma_{xy} & \sigma_y^2
\end{pmatrix}.
\]
As an estimate of \(\theta = \mu_x/\mu_y\), one may use simply \(\theta_n^* = \overline{X}_n/\overline{Y}_n\). Assume \(\mu_y \neq 0\), and find the asymptotic distribution of \(\theta_n^*\).

(b) Specialize to the case where the distribution of \(Y_i\) is exponential with mean 1, and where the conditional distribution of \(X_i\) given \(Y_i = y_i\) is normal with mean, \(\theta y_i\), and variance, 1. First note that \(E(X_i) = \theta\) and \(E(Y_i) = 1\), so that \(\mu_x/\mu_y = \theta\). Find \(\Sigma\) and the asymptotic distribution of \(\theta_n^*\) in this case.

(c) In case (b), find the maximum likelihood estimate of \(\theta\). What is the asymptotic efficiency of \(\theta_n^*\) found in (b).

7. Let \(X_1, \ldots, X_n\) be a sample from the exponential distribution with mean \(\theta\) for some \(\theta > 0\), (density \(f(x|\theta) = \theta^{-1}e^{-x/\theta}I(x > 0)\)), and let \(Y_1, \ldots, Y_n\) be an independent sample from \(f(y|\mu)\) for some \(\mu > 0\).

(a) Find the likelihood ratio test statistic for testing the hypothesis \(H_0: \mu = 2\theta\) based on \(X_1, \ldots, X_n\), and \(Y_1, \ldots, Y_n\).

(b) What is its asymptotic distribution for large \(n\) under \(H_0\)?

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 of symmetry and independence, namely that \(p_{ij} = \pi_i \pi_j\) for some probability vector, \((\pi_1, \ldots, \pi_c)\). 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?
Solutions to the Final Examination, Stat 200C, Spring 2009.

1. (a) $X_n$ converges to $X$ in probability if for all $\epsilon > 0$, $P(|X_n - X| > \epsilon) \to 0$ as $n \to \infty$.
   (b) $X_n$ converges to $X$ in quadratic mean if $E(|X_n - X|^2) \to 0$ as $n \to \infty$.
   (c) This is done using a Chebychev argument. Let $\epsilon > 0$. Since
   \[ E(X_n - X)^2 \geq E((X_n - X)^2I(|X_n - X| > \epsilon)) \geq \epsilon^2 P(|X_n - X| > \epsilon), \]
   we see that if $E(|X_n - X|^2) \to 0$, then $P(|X_n - X| > \epsilon) \to 0$.

2. (a) $E(S_n) = \sum_{k=1}^{n} E(X_k) = \sum_{k=1}^{n} (k/2) = (n+1)/4$.
   $B_n^2 = \text{Var}(S_n) = \sum_{k=1}^{n} \text{Var}(X_k) = \sum_{k=1}^{n} (k^2/12) = (n+1)(2n+1)/72$.
   (b) Let $Z_k = X_k - (k/2)$ so that $EZ_k = 0$ and $|Z_k| < k/2$ a.s. Then
   \[
   \frac{1}{B_n^2} \sum_{1}^{n} E\{Z_k^2I(Z_k^2 > \epsilon^2 B_n^2)\} \leq \frac{1}{B_n^2} \sum_{1}^{n} E\{Z_k^2I(\{(k/2)^2 > \epsilon^2 B_n^2)\} \leq \frac{1}{B_n^2} \sum_{1}^{n} E\{Z_k^2I((n/2)^2 > \epsilon^2 B_n^2)\} = I((n/2)^2 > \epsilon^2 B_n^2)
   \]
   This goes to zero as $n \to \infty$ since $B_n$ goes to infinity at rate $n^3$.

3. (a) The $Z_j$ form a 2-dependent stationary sequence. $E(S_n) = \sum_{1}^{n-2} E(Z_j) = (n-2)E(Z_1) - (n-2)/27$. $\text{Var}(S_n) = (n-2)\text{Var}(Z_1) + 2(n-3)\text{Cov}(Z_1,Z_2) + 2(n-4)\text{Cov}(Z_1,Z_3) = (n-2)(26/27^2) - 2(n-3)/27^2 - 2(n-4)/27^2$.
   (b) $(S_n - ES_n)/\sqrt{n} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 22/27^2)$.

4. (a) The first quartile is $x_{25} = 4/3$ and the median is $x_{5} = 2$.
   (b) $f(x_{25}) = f(4/3) = (3/4)^2$, and $f(x_{5}) = f(2) = (1/2)^2$. So
   \[
   \sqrt{n} \left( \frac{X[\lfloor n/4 \rfloor]}{X[n/2]} - 4/3 \right) \xrightarrow{\mathcal{L}} \mathcal{N}\left( 0, \left( \begin{array}{c} 16/27 \ 
   8/9 \ 
   8/9 \end{array} \right) \right)
   \]
   (c) $P(M_n/b_n \leq x) = P(M_n \leq b_n x) = F(b_n x)^n = (1 - (1/(b_n x)))^n I(x > 1/b_n) \to \lim_{n \to \infty} \exp\{-(n/b_n x)\} I(x > 1/b_n) = \exp\{-1/x\} I(x > 0)$, if we take $b_n = n$.

5. (a) The log likelihood function is
   \[
   \ell_n(\theta) = -\sum_{1}^{n} (X_j - \theta) - 2 \sum_{1}^{n} \log(1 + \exp\{-X_j - \theta\})
   \]
   The likelihood equation is $\hat{\ell}_n = 0$, or
   \[
   n - 2 \sum_{1}^{n} \frac{1}{\exp(X_j - \theta) + 1} = 0.
   \]
(b) Let \( \hat{\theta}_n \) denote the sample median. Then, since \( f(\theta|\theta) = 1/4, \sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{L} \mathcal{N}(0, (1/4)/f(\theta, \theta)^2) = \mathcal{N}(0, 4). \)

(c) We may use \( \hat{\theta}_n = \theta_n - \hat{\ell}_n(\theta_n)^{-1}\hat{\ell}_n(\theta_n) \) or more simply

\[
\hat{\theta}_n = \theta_n + T^{-1}(\theta_n) \frac{1}{n} \hat{\ell}_n(\theta_n) = \theta_n + 3\left[ 1 - \frac{2}{n} \sum_{i=1}^{n} \exp\{X_j - \theta_n\} + 1 \right].
\]

6. (a) By the Central Limit Theorem, \( \sqrt{n}((X_n, Y_n) - (\mu_x, \mu_y)) \xrightarrow{L} \mathcal{N}((0, 0), \varphi). \)

Then we apply Cramér’s Theorem with \( g(x, y) = x/y \) and \( \hat{g}(x, y) = (1/y, -x/y^2) \), so that \( \hat{g}(\mu_x, \mu_y) = (1/\mu_y)(1, -\theta) \). We find

\[
\sqrt{n}(\theta^*_n - \theta) \xrightarrow{L} \mathcal{N}(0, \frac{1}{\mu_y^2}(1, -\theta) \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix} (1, -\theta)) \]

\[
= \mathcal{N}(0, \frac{1}{\mu_y^2}(\sigma_x^2 - 2\theta\sigma_{xy} + \theta^2\sigma_y^2)).
\]

(b) For the exponential distribution of \( Y, \mu_y = 1 \) and \( \sigma_y^2 = 1 \). Then \( E(X) = E(E(X|Y)) = E(\theta Y) = \theta \). Similarly, \( E(X^2) = E(E(X^2|Y)) = E(1 + \theta^2 Y^2) = 1 + 2\theta^2, \)
\( \sigma_x^2 = 1 + \theta^2, \)
\( E(XY) = E(E(XY|Y)) = E(YE(X|Y)) = E(\theta Y^2) = 2\theta, \) and \( \sigma_{xy} = \theta \). So in this case, \( \sqrt{n}(\theta^*_n - \theta) \xrightarrow{L} \mathcal{N}(0, 1). \)

(c) In case (b), the joint density of \((X, Y)\) is \( f(x, y|\theta) = (2\pi)^{-1/2}e^{-y^2}e^{-(x-\theta y)^2/2} \) for \( y > 0 \). We find

\[
(\partial/\partial\theta)\log f(x, y|\theta) = (x - \theta y)y.
\]

From this we can see that the maximum likelihood estimate of \( \theta \) is \( \hat{\theta}_n = (\sum X_i Y_i/\sum Y_i^2) \).

From \( (\partial/\partial\theta)^2 f(x, y|\theta) = -y^2, \) we see that Fisher information is \( I(\theta) = E(Y^2) = 2 \).

Therefore, \( \sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{L} \mathcal{N}(0, 1/2). \) The asymptotic efficiency of \( \theta^*_n \) (relative to \( \hat{\theta}_n \)) is only 50%.

7. (a) \( L_n(\theta, \mu) = \theta^{-n}\mu^{-n}\exp\{-n/\theta \sum X_i - (1/\mu) \sum Y_i \}. \)

The maximum likelihood estimates under the general hypothesis are \( \hat{\theta}_n = \bar{X}_n \) and \( \hat{\mu}_n = \bar{Y}_n \). Under \( H_0 \), they are \( \bar{\theta}_n = (1/2)\bar{X}_n + (1/4)\bar{Y}_n \) and \( \bar{\mu}_n = 2\bar{\theta}_n \). The likelihood ratio test statistic is

\[
\Lambda_n = \frac{L_n(\hat{\theta}_n, \hat{\mu}_n)}{L_n(\hat{\theta}_n, \hat{\mu}_n)} = \frac{2^{-n}\hat{\theta}_n^{-2n}e^{-2n}}{\theta^{-n}\hat{\mu}_n^{-n}e^{-2n} \left[ 2((1/2)\bar{X}_n + (1/4)\bar{Y}_n)^2 \right]^n}
\]

(b) Under \( H_0, -2\log(\Lambda_n) \) converges in law to a chisquare 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 \( \prod_i \prod_j (\pi_i \pi_j)^{n_{ij}} = (\prod_i \pi_i^{n_i}) (\prod_j \pi_j^{n_j}) = \prod_i \pi_i^{n_i} + n_i \). So the maximum likelihood estimates are \( \tilde{\pi}_i = (n_i + n_i) / (2n) \). There are \( c - 1 \) parameters estimated so the chi-square test of \( H_0 \) rejects \( H_0 \) if \( \chi^2(\tilde{\pi}) \) is greater than the appropriate cutoff point for a chi-square distribution with \( (c^2 - 1) - (c - 1) = c(c - 1) \) degrees of freedom.

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