Stat 582
Emerson, Win 12
Homework #1
January 6, 2012

Written problems to be handed in Friday, January 13, 2012.

1. Let $T_n = \sum_{i=1}^n X_i/n$, where $X_1, X_2, \ldots$ are independent random variables having probability distribution given by

$$Pr(X_i = -i) = Pr(X_i = i) = \frac{1}{2}.$$

a. Show that the sequence of statistics $\{T_n\}_{n=1}^\infty$ is asymptotically unbiased for some θ (and identify θ), where a sequence of statistics $\{T_n\}_{n=1}^\infty$ is called *asymptotically unbiased* for θ if

$$\lim_{n\to\infty} E(T_n - \theta) = 0.$$

Ans: Note that $E(X_i) = 0$ for all $i \geq 1$, hence $E(T_n) = 0$ for all $n \geq 1$. Thus T_n is unbiased for $\theta = 0$ and also trivially asymptotically unbiased.

b. Show that the sequence of statistics $\{T_n\}_{n=1}^\infty$ is not consistent for θ.

Ans: The sequence $\{T_n\}_{n=1}^\infty$ is consistent for $\theta = 0$ if

$$\lim_{n\to\infty} Pr(|T_n| > \epsilon) = 0.$$

Hence, we need to show either that no such limit exists, or that it is bounded away from 0.

Note first that for $n \geq 1$,

$$T_{n+1} = \frac{n}{n+1} T_n + \frac{X_{n+1}}{n+1},$$

and because $X_n \in \{-n, n\}$, we have

$$T_{n+1} \in \left\{ \frac{n}{n+1} T_n - 1, \frac{n}{n+1} T_n + 1 \right\}.$$

We thus consider $Pr(|T_{n+1}| > \epsilon)$ using

$$Pr(|T_{n+1}| > \epsilon) = Pr(|T_{n+1}| > \epsilon \,\cap\, T_n \leq -\epsilon) + Pr(|T_{n+1}| > \epsilon \,\cap\, |T_n| < \epsilon) + Pr(|T_{n+1}| > \epsilon \,\cap\, T_n \geq \epsilon).$$

Notationally, we define

$$Pr(|T_n| < \epsilon) = p_{n,\epsilon}$$

and note that the symmetry of distribution of the X_is then provides that

$$Pr(T_n < -\epsilon) = Pr(T_n > \epsilon) = \frac{(1 - p_{n,\epsilon})}{2}.$$

Now, we note that

$$Pr(|T_{n+1}| > \epsilon \,\cap\, T_n \leq -\epsilon) \geq Pr(X_{n+1} = -(n+1) \,\cap\, T_n \leq -\epsilon) = \frac{1}{2} Pr(T_n \leq -\epsilon) = \frac{(1 - p_{n,\epsilon})}{4}$$
$$Pr(|T_{n+1}| > \epsilon \,\cap\, T_n \geq \epsilon) \geq Pr(X_{n+1} = (n+1) \,\cap\, T_n \geq \epsilon) = \frac{1}{2} Pr(T_n \geq \epsilon) = \frac{(1 - p_{n,\epsilon})}{4}$$

.

Now suppose that $|T_n| < \epsilon$. Then, considering the cases based on the value of X_{n+1} we have

$$-\epsilon < T_n < \epsilon \qquad \Rightarrow \qquad \begin{cases} -\frac{n}{n+1}\epsilon - 1 < T_{n+1} < \frac{n}{n+1}\epsilon - 1 & \text{if } X_{n+1} = -(n+1), \text{ and} \\ -\frac{n}{n+1}\epsilon + 1 < T_{n+1} < \frac{n}{n+1}\epsilon + 1 & \text{if } X_{n+1} = n+1, \end{cases}$$

which for $\epsilon < (n+1)/n$ yields

$$|T_{n+1}| > 1 - \frac{n}{n+1}\epsilon.$$

Furthermore, because

$$\epsilon < \frac{(n+1)}{(2n+1)} \qquad \Rightarrow 1 - \frac{n}{n+1}\epsilon > \epsilon$$

and $(n+1)/(2n+1)$ is decreasing in n, we know that for any $0 < \epsilon < \lim_{n\to\infty} \frac{n+1}{2n+1} = \frac{1}{2}$,

$$|T_n| < \epsilon \qquad \Rightarrow \qquad |T_{n+1}| > \epsilon,$$

and for $\epsilon < 0.5$,

$$Pr(|T_{n+1}| > \epsilon \,\cap\, |T_n| < \epsilon) = Pr(|T_n| < \epsilon) = p_{n,\epsilon}.$$

Combining the parts, we thus find for $0 < \epsilon < 0.5$ and $n > 0$

$$Pr(|T_{n+1}| > \epsilon) \geq \frac{(1 - p_{n,\epsilon})}{4} + p_{n,\epsilon} + \frac{(1 - p_{n,\epsilon})}{4} = \frac{(1 + p_{n,\epsilon})}{2} > \frac{1}{2},$$

and $T_n \not\to_p 0$.

2. Chebyshev's inequality states that for a random variable X having finite variance σ^2 and for any $\epsilon > 0$

$$Pr\left(|X - E(X)| \geq \epsilon\sigma\right) \leq \frac{1}{\epsilon^2}.$$

a. Show that the upper bound on the tail probabilities can be extremely conservative by finding a distribution for which the true bound is 0 for all $\epsilon > 1$.

Ans: Chebyshev's inequality places bounds on the probability that a random variable X might be more than ϵ standard deviations away from its mean. This problem asks you to find a distribution in which the entire support of the distribution is within 1 standard deviation of the mean. Clearly, when σ is finite, this is only possible for variables having a distribution with finite support. For a random variable X with a distribution having support (a, b), mean $\mu < \infty$, and variance $\sigma^2 < \infty$, we then want to consider whether

$$max\left[\frac{|a - \mu|}{\sigma}, \frac{|b - \mu|}{\sigma}\right] = 1,$$

because in that case the entire support of X is within one standard deviation of the mean and the probability of being more than one standard deviation from the mean (corresponding to a choice of $\epsilon > 1$ in Chebyshev's inequality) is less than $1/\epsilon^2$ for

Obvious families of distributions to consider are

- Beta distributions: $X \sim Beta(\alpha > 0, \beta > 0)$ has a wide variety of shapes for probability density functions with skewness ranging from $-\infty$ to ∞. Choices of $\alpha = \beta = 0.5$ correspond to the uniform $\mathcal{U}(0, 1)$ distribution, and choices of $\alpha = \beta$ close to zero will approach a Bernoulli distribution with $p = 0.5$. Properties of the beta distribution include $X \in (0, 1)$, $\mu \equiv EX = \alpha/(\alpha + \beta)$, and

$\sigma^2 \equiv Var(X) = \alpha\beta/\left[(\alpha+\beta)^2(\alpha+\beta+1)\right]$. We are thus interested in whether $\epsilon = 1$ (and the values of α and β at $\epsilon = 1$) where

$$\epsilon = \inf_{\alpha>0,\beta>0} \max\left[\sqrt{\frac{\alpha(\alpha+\beta+1)}{\beta}}, \sqrt{\frac{\beta(\alpha+\beta+1)}{\alpha}}\right].$$

Noting the symmetry in α and β for the two terms inside the max[ ], and that the second term is at least as large as the first when $\alpha \le \beta$, we then want to determine whether $\epsilon = 1$ where

$$\epsilon = \inf_{\alpha>0,\beta>\alpha} \left\{\sqrt{\frac{\beta(\alpha+\beta+1)}{\alpha}}\right\}.$$

Now for each β, ϵ is clearly decreasing in α, so ϵ will be smallest when $\alpha = \beta$, in which case $\epsilon = \inf_{\beta>0}\sqrt{2\beta+1}$, which is 1 as $\beta \to 0$. So a beta distribution with very small $\alpha = \beta$ approaches our criterion, but cannot meet it exactly.

- Bernoulli distributions: $X \sim Bernoulli(p) = \mathcal{B}(1,p)$ has a family of distributions for which the skewness can range from $-\infty$ to ∞ as p ranges from near 1 down to near 0. Properties of the Bernoulli include $X \in \{0,1\}$, $\mu \equiv EX = p$, and $\sigma^2 \equiv Var(X) = p(1-p)$. We are thus interested in whether $\epsilon = 1$ (and the value of p at $\epsilon = 1$) where

$$\epsilon = \inf_{p\in(0,1)} \max\left[\sqrt{\frac{p}{1-p}}, \sqrt{\frac{1-p}{p}}\right].$$

Noting the symmetry about $p = 0.5$ for the two terms insidemax[ ] and that the second term is at least as large as the first when $p \le 0.5$, we then want to determine whether $\epsilon = 1$ where

$$\epsilon = \inf_{p\in(0,0.5)} \left\{\sqrt{\frac{1-p}{p}}\right\}.$$

Noting that $(1-p)/p$ is decreasing in p, the smallest value will be where $p = 0.5$ and $\epsilon = 1$, meeting our criterion.

b. Show that the upper bound is not always conservative by showing that for each choice of $\epsilon > 1$ there is some random variable X_ϵ that has the tail probabilities exactly equal to the Chebyshev bound.

Ans: This problem is asking you to show that for any choice of $\epsilon > 1$, there is some distribution for which the probability of being at least ϵ standard deviations away from the mean is $1/\epsilon^2$. Having succeeded in the previous problem with a discrete distribution over finite support, we again consider such distributions.

- For a Bernoulli distribution with $X \sim \mathcal{B}(1,p)$, $|X-p| \in \{p, 1-p\}$. For $\epsilon > 1$, we need the probability of being at least ϵ standard deviations away from the mean to be less than 1, so we need to focus on cases in which, say, 0 is less than ϵ standard deviations of the mean, and 1 is exactly ϵ standard deviations away from the mean. (We could of course have reversed the roles of 0 and 1 for this distribution, owing to the symmetry of the variance $p(1-p)$ about $p = 0.5$.) Hence we want

$$\frac{1}{\epsilon^2} = p \qquad \text{and} \qquad \frac{1-p}{\sqrt{p(1-p)}} = \sqrt{\frac{1-p}{p}} = \epsilon,$$

which simultaneous equations have no solution.

- We instead try a trinomial distribution where $Pr(X = 1) = Pr(X = -1) = p$ and $Pr(X = 0) = 1 - 2p$ for $0 < p \le 0.5$. This is a symmetric distribution, so the skewness is always 0. Other properties include $\mu \equiv EX = 0$, $\sigma^2 \equiv Var(X) = 2p$, kurtosis $\mu_4 \equiv E[X-0]^4 = 2p$, and coefficient of excess kurtosis $\gamma_2 \equiv \mu_u/\sigma^4 - 3 = 1/(2p) - 3$. Note that within this family, the coefficient of excess kurtosis is smallest for $p = 0.5$ (in which case $\gamma_2 = -2$) and becomes infinite as $p \to 0$. As the mean is 0 for each $p \in [0, 0.5]$, we just need to consider when an observation $X \in \{-1, 1\}$ is ϵ standard deviations away from 0. Hence we want

$$\frac{1}{\epsilon^2} = 2p \qquad \text{and} \qquad \frac{1-0}{\sqrt{2p)}} = \epsilon,$$

which two equations are identical. So for any $\epsilon \geq 1$, this trinomial distribution with $p = \epsilon/2$ exactly achieves the Chebyshev bound. Note that when the Chebyshev bound is exact for a given $\epsilon = \epsilon_0$, the Chebyshev bound is conservative for all $\epsilon \neq \epsilon_0$, with the probability of being more than ϵ standard deviations away from the mean being 0 for $\epsilon > \epsilon_0$. (Note also that for $\epsilon = 1$, we would choose $p = 0.5$, and $(X+1)/2 \sim calB(1, p)$, so this "trinomial" distribution degenerates to the Bernoulli distribution we found in part a.)

3. Let sequence of statistics $\{T_n\}_{n=1}^{\infty}$ be asymptotically unbiased for θ. Furthermore, suppose that

$$\lim_{n\to\infty} Var(T_n) \to 0.$$

Show that $\{T_n\}_{n=1}^{\infty}$ is consistent for θ.

Ans: Let $\mu_n = E(T_n)$ and $b_n = \mu_n - \theta$ be the bias of T_n as an estimator of θ. Let $\sigma_n^2 = Var(T_n)$. By supposition, $b_n \to 0$ and $\sigma_n^2 \to 0$. Now because $T_n - \theta = T_n - \mu_n + \mu_n - \theta = T_n - \mu_n + b_n$, we have for $\epsilon > 0$

$$\begin{aligned} Pr(|T_n - \theta| \geq \epsilon) &= 1 - Pr(-\epsilon < T_n - \theta < \epsilon) \\ &= 1 - Pr(-\epsilon - b_n < T_n - \mu_n < \epsilon - b_n) \\ &\leq 1 - Pr(|T_n - \mu_n| < \epsilon + |b_n|) \\ &= Pr(|T_n - \mu_n| \geq \epsilon + |b_n|). \end{aligned}$$

so by Chebyshev's inequality, for all $\epsilon > 0$

$$Pr(|T_n - \theta| \geq \epsilon) \leq Pr(|T_n - \mu_n| \geq \epsilon + |b_n|) \leq \frac{\sigma_n^2}{(\epsilon + |b_n|)^2} \to 0,$$

so $T_n \to_p \theta$.

4. Let $X_1, X_2, \ldots$ be independent random variables for which $X_n \sim \chi_n^2$, a chi squared distribution with n degrees of freedom. Let sequence of statistics $\{T_n\}_{n=1}^{\infty}$ be defined by $T_n = X_n/n$. Show that $\{T_n\}_{n=1}^{\infty}$ is consistent for some θ (and identify θ).

Ans: Let $Z_1, Z_2, \ldots$ be independent, identically distributed random variables with $Z_i \sim \mathcal{N}(0, 1)$. Then, by the definition of the chi squared distribution, for $Y_i = Z_i^2$, we have $Y_i \sim \chi_1^2$, with $EY_i = 1$ and $Var(Y_i) = 2$, as well as $\sum_{i=1}^{n} Y_i \sim \chi_n^2$, a chi squared distribution with n degrees of freedom. Thus for

$$\overline{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

we have $n\overline{Y}_n \sim \chi_n^2$, and $n\overline{Y}_n \sim X_n$, hence $T_n = X_n/n \sim \overline{Y}_n$. But by the weak law of large numbers, $\overline{Y}_n \to_p 1$, so we have $T_n \to_p \theta = 1$.

5. Suppose sequence of random variables $\{X_n\}_{n=1}^{\infty}$ converge almost surely to random variable X. Suppose further that function g is continuous. Rigorously prove that $g(X_n) \to_{as} g(X)$.

Ans: $X_n \to_{as} X$ dictates that $P(A) = 1$ for the set

$$A = \{\omega : X_n(\omega) \to X(\omega)\}.$$

We will show that on that $\forall \omega \in A$, $g(X_n(\omega)) \to g(X(\omega))$.

Let $\omega \in A$ be arbitrary, and define $x_n \equiv X_n(\omega)$ and $x \equiv X(\omega)$, we have the desired convergence if and only if

$$\forall \epsilon > 0 \ \exists n_\epsilon \quad \text{such that} \quad \forall n > n_\epsilon \ |g(x_n) - g(x)| < \epsilon.$$

Now for arbitrary $\epsilon > 0$, because $g(\cdot)$ is continuous, we know that

$$\exists \delta_\epsilon \quad \text{such that} \quad |x^* - x| < \delta_\epsilon \ \Rightarrow |g(x^*) - g(x)| < \epsilon,$$

so if we know $|x_n - x| < \delta_\epsilon$ for sufficiently large n, we have our result. But this is known because $x_n \to x$, which provides that

$$\exists n_{\delta_\epsilon} \quad \text{such that} \quad \forall n > n_{\delta_\epsilon}, \ |x_n - x| < \delta_\epsilon.$$

Since this holds for arbitrary $\omega \in A$, we have $g(X_n) \to_{as} g(X)$.

6. Provide counterexamples demonstrating the falseness of each of the following statements.

Ans: I tried to choose as trivial examples as possible. Hence, I often chose situations in which the X_n's are i.i.d., if not identically equal to each other. In all cases, I consider the probability space defined with Lebesgue measure on the Borel sets on the unit interval: $(\Omega = [0,1], \mathcal{B}(\Omega), \lambda)$.

a. $X_n \to_d X$ implies $X_n \to_p X$.

Ans: Let $Y(\omega) = 1_{[0,0.5]}(\omega)$. Then $Y \sim \mathcal{B}(1, 0.5)$. Then $Y \sim 1 - Y$, so letting $X_n = Y$ for $n \geq 1$, and letting $X = 1 - Y$, we have $X_n \to_d X$, but $Pr(X_n = X) = 0$ for $n \geq 1$, so $Pr(\{\omega : X_n(\omega) \to X(\omega)\}) = 0 \not\to 1$.

b. $X_n \to_d X$ implies $X_n \to_{as} X$.

Ans: Same counterexample as in part a shows this as well.

c. $X_n \to_p X$ implies $X_n \to_{as} X$.

Ans: Define indicator random variables $Y_{m,j}$ for $m = 1, 2, \ldots$ and $j = 1, \ldots, m$ as

$$Y_{m,j} = \begin{cases} 1 & \omega \in \left(\frac{j-1}{m}, \frac{j}{m}\right] \\ 0 & \text{else} \end{cases} .$$

Then for $m = 1, 2, \ldots$, $j = 1, \ldots, m$, and $n = m(m-1)/2 + j$, define $X_n = Y_{m,j}$ (note that this definition does describe a 1:1 correspondence between the X_n's and the $Y_{m.j}$'s, because $m(m-1)/2 + m + 1 = (m+1)m/2 + 1$.). Thus for m and j corresponding to some n

$$Pr(\{\omega : X_n(\omega) \neq 0\}) = Pr(\{\omega : X_n(\omega) = 1\}) = Pr\left[\omega \in \left(\frac{j-1}{m}, \frac{j}{m}\right]\right) = \frac{1}{m} \to 0 \text{ as } n \to \infty.$$

Hence $X_n \to_p X = 0$. However

$$\forall \omega \in (0,1] \ \forall n \ \exists n_0, n_1 > n \text{ such that } X_{n_0} = 0 \text{ and } X_{n_1} = 1.$$

Hence

$$Pr(\{\omega : X_n(\omega) \to X(\omega) = 0\}) = Pr(\omega = 0) = 0 \neq 1,$$

so $X_n \not\to_{as} 0$.

d. $X_n \to_p X$ implies $X_n \to_r X$ for $r = 1$.

Ans: Using the indicator random variables $Y_{m,j}$ defined in part c, now let $X_n = nY_{n,n}$. Note that $X_n \sim n\mathcal{B}(1, p = 1/n)$, so $X_n \sim (\mu = 1, \sigma^2 = n - 1)$. We thus have

$$Pr\left(\{\omega : X_n(\omega) \neq 0\}\right) = Pr\left[\omega \in \left(\frac{n-1}{n}, 1\right]\right) = \frac{1}{n} \to 0 \text{ as } n \to \infty.$$

Hence $X_n \to_p X = 0$. However, $E[|X_n - X|] = E[|X_n|] = 1 \not\to 0$ as $n \to \infty$ so $X_n \not\to_{r=1} X = 0$. (Note that this sequence of random variables is not uniformly integrable.)

e. $X_n \to_r X$ implies $X_n \to_{r'} X$ for $r' > r$.

Ans: Using the indicator random variables $Y_{m,j}$ defined in part c, now let $X_n = \sqrt{n}Y_{n,n}$. Note that $X_n \sim \sqrt{n}\mathcal{B}(1, p = 1/n)$, so $X_n \sim (\mu = 1\sqrt{n}, \sigma^2 = (n-1)/n)$. We thus have

$$Pr\left(\{\omega : X_n(\omega) \neq 0\}\right) = Pr\left[\omega \in \left(\frac{n-1}{n}, 1\right]\right) = \frac{1}{n} \to 0 \text{ as } n \to \infty.$$

Hence $X_n \to_p X = 0$. We also have, $E[|X_n - X|] = E[|X_n|] = 1/\sqrt{n} \to 0$ as $n \to \infty$ so $X_n \to_{r=1} X = 0$. However, $E[|X_n - X|^2] = E[|X_n|^2] = \frac{(n-1)}{n} + \frac{1}{n} = 1 \not\to 0$ as $n \to \infty$ so $X_n \not\to_{r=2} X = 0$. (Note that the sequence of random variables X_n^2 is not uniformly integrable.)

7. Let $X_1, X_2, \ldots$ be random variables having means $EX_i = \mu_i < \infty$, variance $Var(X_i) = \sigma^2 < \infty$, with $corr(X_i, X_j) = 0$ for $i \neq j$ (note that the variables need not be independent, just uncorrelated). Define $\overline{X}_n = \sum_{i=1}^n X_i/n$ and $\theta_n = \sum_{i=1}^n \mu_i/n$. Show that the sequence of statistics $\{T_n\}_{n=1}^\infty$ with $T_n = \overline{X}_n - \theta_n$ satisfies $T_n \to_p 0$.

Ans: Define $W_n = X_n - \mu_n$, and note that $\overline{W}_n = T_n$, that $W_n \sim (0, \sigma^2)$, and that $corr(W_i, W_j) = corr(X_i, X_j)$, so $cov(W_i, W_j) = \sigma^2 1_{[i=j]}$. Then by linearity of expectation,

$$\begin{aligned} Var(T_n) = E\left[\sum_{i=1}^n \frac{W_i}{n}\right]^2 &= E\left[\sum_{i=1}^n \sum_{j=1}^n \frac{W_i W_j}{n^2}\right] \\ &= \sum_{i=1}^n \sum_{j=1}^n \frac{E[W_i W_j]}{n^2} \\ &= \sum_{i=1}^n \sum_{j=1}^n \frac{cov[W_i W_j]}{n^2} \\ &= \sum_{i=1}^n \frac{\sigma^2}{n^2} = \frac{\sigma^2}{n}. \end{aligned}$$

Hence $T_n \sim (0, \sigma^2/n)$, and by Chebyshev's inequality

$$Pr(|T_n - 0| \geq \epsilon) \leq \frac{Var(T_n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \to 0$$

as $n \to \infty$. Hence, $T_n \to_p 0$.