Kris Harper
STAT 24400
November 11, 2010

Homework 6

Problem 1. *Let X_i be as in Problem 5.1 but with $E(X_i) = \mu_i$ and $n^{-1}\sum_{i=1}^n \mu_i \to \mu$. Show that $\overline{X} \to \mu$ in probability.*

Proof. By Chebyshev's inequality we have

$$P(|X_i - \mu_i| > \varepsilon) \leq \frac{\mathrm{Var}(X_i)}{\varepsilon^2}$$

for each i and therefore we have

$$P\left(\left|\frac{1}{n}\sum_{i=1}^n X_i - E\left(\frac{1}{n}\sum_{i=1}^n X_i\right)\right| > \varepsilon\right) \leq \frac{1}{\varepsilon^2}\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^n X_i\right).$$

Reducing both sides and noting $E(X_i) = \mu_i$, we have

$$P\left(\left|\overline{X}_n - \frac{1}{n}\sum_{i=1}^n \mu_i\right| > \varepsilon\right) \leq \frac{1}{(\varepsilon n)^2}\sum_{i=1}^n \mathrm{Var}(X_i).$$

From Problem 5.1, $n^{-2}\sum_{i=1}^n \mathrm{Var}(X_i) \to \sigma^2$, and we know $n^{-1}\sum_{i=1}^n \mu_i \to \mu$. Thus, as $n \to \infty$, both sides of this inequality reduce to

$$P(|\overline{X}_n - \mu| > \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$$

and the righthand side goes to 0. Thus $\overline{X} \to \mu$ in probability. □

Problem 2. *Suppose that the number of insurance claims, N, filed in a year is Poisson distributed with $E(N) = 10,000$. Use the normal approximation to the poisson to approximate $P(N > 10,200)$.*

We use standardization so we have

$$P(N > 10200) = P\left(\frac{N - 10000}{\sqrt{10000}} > \frac{10200 - 10000}{\sqrt{10000}}\right) \approx 1 - \Phi(2) = .0228$$

Problem 3. *Show that if $X_n \to c$ in probability and if g is a continuous function, then $g(X_n) \to g(c)$ in probability.*

Proof. Since $X_n \to c$ in probability we know for each $\varepsilon > 0$ $P(|X_n - c| > \varepsilon) \to 0$. Since g is continuous we know for each $\eta > 0$ there is some $\delta > 0$ such that for each X_i we have $|X_n - c| < \delta$ implies $|g(X_n) - g(c)| < \eta$. Thus for each X_i, we see that $|g(X_i) - g(c)|$ is bounded given that $|X_i - c|$ can be bounded. Since $P(|X_n - c| > \varepsilon) \to 0$, we then know that $P(|g(X_n) - g(c)| > \varepsilon) \to 0$ as well. □

Problem 4. *A skeptic gives the following argument to show that there must be a flaw in the central limit theorem: "We know that the sum of independent Poisson random variables follows a Poisson distribution with a parameter that is the sum of the parameters of the summands. In particular, if n independent Poisson random variables, each with parameter n^{-1}, are summed, the sum has a Poisson distribution with parameter 1. The central limit theorem says that as n approaches infinity, the distribution of the sum tends to be a normal distribution, but the Poisson with parameter 1 is not the normal." What do you think of this argument?*

The argument can be stated as follows. Fix n, and let $X_1, \ldots, X_n$ be n independent Poisson random variables with parameter n^{-1}. As it's stated, the central limit theorem doesn't apply because we need an infinite sequence of random variables, but we only have n. We can try to increase the number of variables n, but to keep the conditions on the parameter, we must also decrease the parameter n^{-1}. As $n \to \infty$, $n^{-1} \to 0$ so once we have a sequence of such variables, they have parameter 0, which is no longer a Poisson distribution.

Essentially, this argument is invalid because you can't vary your parameter as you increase your sequence of variables. You must start with a sequence of variables with fixed parameters.

Problem 5. *Suppose that $X_1, \ldots, X_{20}$ are independent random variables with density functions*

$$f(x) = 2x, \quad 0 \le x \le 1.$$

Let $S = X_1 + \cdots + X_{20}$. Use the central limit theorem to approximate $P(S \le 10)$.

From the last homework we know $E(X_i) = \frac{2}{3}$ and $\mathrm{Var}(X_i) = \frac{1}{18}$. Then $E(S - 40/3) = E(S) - 40/3 = 20E(X_i) - 40/3 = 0$ and $\mathrm{Var}(S - 40/3) = \mathrm{Var}(S) = 20/18$. Then the central limit theorem says

$$\begin{aligned}
P(S \le 10) &= P\left(S - \frac{40}{3} \le 10 - \frac{40}{3}\right) \\
&= P\left(\frac{S - \frac{40}{3}}{\sqrt{\frac{20}{18}}\sqrt{20}} \le \frac{10 - \frac{40}{3}}{\sqrt{\frac{20}{18}}\sqrt{20}}\right) \\
&\approx \Phi\left(\frac{10 - \frac{40}{3}}{\sqrt{\frac{20}{18}}\sqrt{20}}\right) \\
&= \Phi\left(\frac{-1}{\sqrt{2}}\right) \\
&\approx \Phi(-.707) \\
&\approx .2206.
\end{aligned}$$

Problem 6. *Suppose that a company ships packages that are variable in weight, with an average weight of* 15 *lb and a standard deviation of* 10. *Assuming that the packages come from a large number of different customers so that it is reasonable to model their weights as independent random variables, find the probability that* 100 *packages will have a total weight exceeding* 1700 *lb.*

Let X_i be the weight of a given package so that $E(X_i) = 15$ and $\mathrm{Var}(X_i) = 100$. Let $S_n = \sum_{i=1}^n X_i$. We want to find $P(S_n \ge 1700)$. By the central limit theorem this is

$$\begin{aligned}
P(S_n - 100 \cdot 15 \ge 1700 - 100 \cdot 15) &= P\left(\frac{S_n - 100 \cdot 15}{10 \cdot \sqrt{100}} \ge \frac{1700 - 100 \cdot 15}{10 \cdot \sqrt{100}}\right) \\
&= 1 - P\left(\frac{S_n - 1500}{100} \le \frac{200}{100}\right) \\
&\approx 1 - \Phi\left(\frac{200}{100}\right) \\
&= 1 - \Phi(2) \\
&\approx .0228.
\end{aligned}$$

Problem 7. *Suppose that X is a discrete random variable with*

$$P(X = 0) = \frac{2}{3}\theta$$

$$P(X=1) = \frac{1}{3}\theta$$

$$P(X=2) = \frac{2}{3}(1-\theta)$$

$$P(X=3) = \frac{1}{3}(1-\theta)$$

where $0 \le \theta \le 1$ *is a parameter. The following* 10 *independent observations were taken from such a distribution:* $(3, 0, 2, 1, 3, 2, 1, 0, 2, 1)$.
(e) If the prior distribution of Θ *is uniform on* $[0,1]$*, what is the posterior density? Plot it. What is the mode of the posterior?*

We can write the prior density as

$$f_{X_i|\Theta}(x_i \mid \theta) = \begin{cases} \frac{2}{3}\theta & x_i = 0 \\ \frac{1}{3}\theta & x_i = 1 \\ \frac{2}{3}\theta & x_i = 2 \\ \frac{1}{3}\theta & x_i = 3 \end{cases}.$$

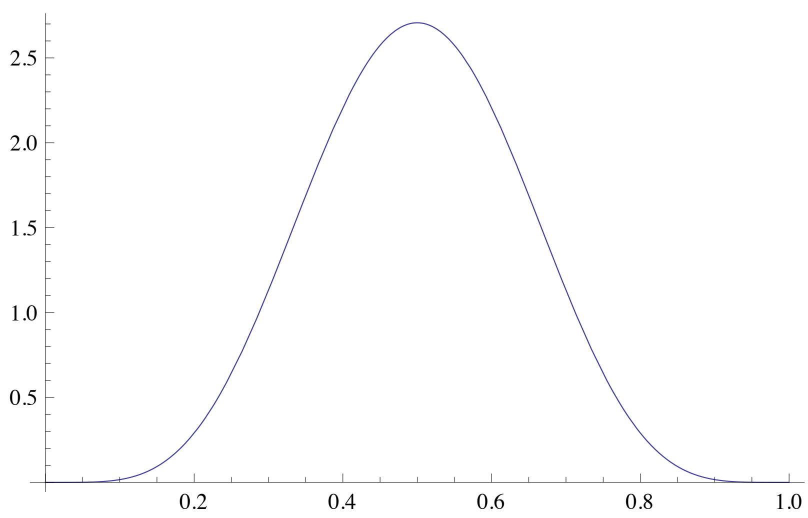
The total joint density for the experiment is given (by independence) as the product of each of the marginal densities for each X_i. Since we know the numbers of each possible instance of x_i, we can now write this product as

$$f_{X|\Theta}(x \mid \theta) = \left(\frac{2}{3}\theta\right)^2 \left(\frac{1}{3}\theta\right)^3 \left(\frac{2}{3}(1-\theta)\right)^3 \left(\frac{1}{3}(1-\theta)\right)^2 = \frac{32}{59049}(1-\theta)^5\theta^5.$$

The posterior density is given by

$$f_{\Theta|X}(\theta \mid x) = \frac{f_{X|\Theta}(x \mid \theta) f_\Theta(\theta)}{\int f_{X|\Theta}(x \mid \theta) f_\theta(\theta) d\theta} = \frac{(1-\theta)^5\theta^5}{\int_0^1 (1-\theta)^5\theta^5 d\theta} = 2772(1-\theta)^5\theta^5 = \frac{\Gamma(6+6)}{\Gamma(6)\Gamma(6)}(1-\theta)^5\theta^5$$

so this is a beta distribution with parameters 6 and 6. The following is a plot of $f_{\Theta|X}$.



To find the mode we must maximize $f_{\Theta|X}(\theta \mid x) = 2772(1-\theta)^5\theta^5$. Taking the derivative we have

$$f'_{\Theta|X}(\theta \mid x) = 13860((1-\theta)^5\theta^4 - (1-\theta)^4\theta^5)$$

and setting this equal to zero gives the maximum at $\theta = 1/2$.

Problem 8. *Suppose that X follows a geometric distribution,*

$$P(X = k) = p(1-p)^{k-1}$$

and assume an i.i.d. sample of size n.
(d) Let p have a uniform prior distribution on $[0,1]$. What is the posterior distribution of p? What is the posterior mean?

Let $X_1, \ldots, X_n$ be the n i.i.d. observations. Then for an arbitrary X_i, we have the distribution

$$f_{X_i|p}(x_i \mid p) = p(1-p)^{x_i - 1}.$$

By independence, the joint distribution is the product of the marginals

$$f_{X|p}(x \mid p) = p^n(1-p)^{\sum_{i=1}^n x_i - n}.$$

The posterior distribution is then given by

$$\begin{aligned} f_{p|X}(p \mid x) &= \frac{(1-p)^{\sum_{i=1}^n x_i - n} f_p(p)}{\int (1-p)^{\sum_{i=1}^n x_i - n} f_p(p) dp} \\ &= \frac{(1-p)^{\sum_{i=1}^n x_i - n}}{\int_0^1 (1-p)^{\sum_{i=1}^n x_i - n} dp} \end{aligned}$$

where $0 \leq p \leq 1$. But now note that the denominator is $B\left(1, 1-n+\sum_{i=1}^n x_i\right)$, where $B(\alpha, \beta)$ is the beta function. Thus, the $f_{p|X}$ is a beta distribution with parameters 1 and $1 - n + \sum_{i=1}^n x_i$.

The posterior mean for a beta distribution with parameters α and β is given by $\alpha/(\alpha+\beta)$. So in our case

$$\mu = \frac{1}{2 - n + \sum_{i=1}^n x_i}.$$