dc3451
David Chen

Assignment 6

6.2.1

Note that $\lim_{n\to\infty} \mu_n = 0 \implies \exists N \mid \forall n > N, |\mu_n| < \epsilon$ for any positive ϵ. Then, we have that $|X_n - 0| > \epsilon \iff X_n > \epsilon$, and the Markov inequality yields that

$$P(X_n \geq \epsilon) \leq \frac{\mu_n}{\epsilon}$$

Taking the respective N such that $\mu_n < \epsilon^2$ for $n > N$, we have the desired result.

6.2.5

The Chebyshev inequality yields that

$$P(|X - \mu| \leq 2\sigma) \geq 1 - \frac{1}{4n} \geq 0.99$$

Thus, $n \geq 25$

6.2.15

g is continuous allows that $\lim_{z\to b} g(z) = g(b)$, such that for any $\epsilon \exists \delta$ that on the δ-ball around b we have $|g(z) - g(b)| < \epsilon$.

We have that $Z_n \to b \implies P(|Z_n - b| < \epsilon) = 1$; taking this $\epsilon = \delta$ from before, we get that $P(|Z_n - b| < \delta) \leq P(|g(Z_N) - g(b)| < \epsilon)$, and as the left approaches 1 in the limit, the right hand does as well.

6.3.4

X_n ought to be roughly normal, with mean μ and variance $9/n$. Then, $Z = \sqrt{n}(X - \mu)/3$ is the standard normal distribution, such that

$$P(|X_n - \mu| < 0.3) = P(|3Z/\sqrt{n}| < 0.3) = P(|Z| < 0.1\sqrt{n}) = 2\Phi(0.1\sqrt{n}) - 1$$

Then, you need $n \geq 384.2 \implies n = 385$.

6.3.10

a

$$P(|X_n - \mu| \geq \frac{\sigma}{4}) \leq \frac{\sigma^2}{n\frac{\sigma^2}{4^2}} = \frac{16}{n} \implies P(|X_n - \mu| \leq \frac{\sigma}{4}) \geq 1 - \frac{16}{n}$$

Here, we need $1 - 16/n \geq 0.99 \implies n \geq 1600$.

b

X_n ought to be distributed approximately normally, such that the standard normal distribution $Z = \sqrt{n}(X_n - \mu)/\sigma$

$$P(|X_n - \mu| \leq \frac{\sigma}{4}) = P(|Z| \leq \frac{\sqrt{n}}{4}) = 2\Phi(\frac{\sqrt{n}}{4}) - 1$$

Here, we need $n \geq 105.4 \implies n = 106$, which is a significantly better bound.

6.3.12

The mgf of the binomial distribution here with parameters n, p_n is $\psi_n(t) = (p_n e^t + 1 - p_n)^n$.

$$\lim_{n\to\infty} \psi_n(t) = \lim_{n\to\infty} (1 - (1 - e^t)p_n)^n = \lim_{n\to\infty} (1 - (1 - e^t)\frac{p_n n}{n})^n = e^{\lambda[e^t - 1]}$$

which is the desired mgf.

6.5.11

a

Note that the gamma distributions is the distribution of the sum of n independent and identical exponential random variables with parameter 3. Thus, if n is large, then the value of that above sum divided by n ought to be normal, and thus the distribution of the sum ought to be normal.

b

The mean and variance of each such exponential distribution is $\frac{1}{3}, \frac{1}{9}$. Thus, we have from the CLT that their averages ought to be normally distributed with mean $\frac{1}{3}$ and variance $\frac{1}{9n}$. Then, the overall sum ought to be normal with mean $\frac{n}{3}$ and variance $\frac{n}{9}$.

6.5.12

a

Note that the negative binomial distribution is the sum of n independent and identical geometric random variables. Thus, if n is large, then the value of the sum divided by n ought to be normal, and thus the sum itself is normal.

b

The mean and variance of each geometric distribution is 4 and 20, and as above we have the sum then must be normal with mean and variance $4n$ and $20n$.

7.1.1

We have the observable variables X_i and one parameter P. In that case, each X_i is Bernoulli with parameter p given $P = p$ and are independent of each other.

7.1.6

We have that the random observable variables are X, the amount of Mexican-American jurors, and the hypothetically observable P which is the proportion of Mexican-Americans among all grand jurors.

Note that P has a beta distribution which has unspecified parameters, and the conditional distribution of X given $P = p$ is the binomial distribution with parameters 220 and p.

7.2.2

We have that the joint pf of $x = (X_1, \ldots, X_n)$ is

$$f(x \mid \theta) = \theta^2(1-\theta)^6$$

Then, we have from Bayes that

$$\xi(\theta \mid x) = \frac{f(x \mid \theta)\xi(\theta)}{f(x \mid \theta_1)\xi(\theta_1) + f(x \mid \theta_2)\xi(\theta_2)}$$

so for $\theta = 0.1$,

$$\xi(0.1 \mid x) = \frac{\xi(0.1)(0.1)^2(0.9)^6}{\xi(0.1)(0.1)^2(0.9)^6 + \xi(0.2)(0.2)^2(0.8)^2} = 0.542$$

Then, $\xi(0.2 \mid x) = 1 - 0.542 = 0.458$.

7.2.6

This forms a posterior beta distribution with parameters $\alpha = 3 + 1 = 4$, $\beta = 5 + 1 = 6$, as this is exactly the construction of such a distribution in many given examples.

7.3.1

The posterior means is

$$\frac{20v^2(0.125)}{100+20v^2} = 0.12 \implies v^2 = 120$$

7.3.2

We have y defectives and z nondefectives, and

$$V = \frac{(y+1)(z+1)}{(y+z+2)^2(y+z+3)}$$

Put $n = y+z$, such that

$$V = \frac{(y+1)((n-y)+1)}{(n+2)^2(n+3)}$$

This is maximized by taking $y = \frac{n}{2}$, rounded up (or down). Then, if n is even, we have

$$V = \frac{(\frac{n}{2}+1)^2}{(n+2)^2(n+3)}$$

Taking $n = 22$, we have $V = 0.01$ exactly, and so for $n > 22$, $V < 0.01$.
This is minimized when $y = n$ or $y = 0$, such that

$$V = \frac{n+1}{(n+2)^2(n+3)}$$

We see that for $n = 7$, $V = 0.0098$, and $n = 6$, $V = 0.012$ and so $V < 7$ must be greater than 0.01.

7.3.13

For a gamma distribution with parameters α, β, we have that

$$\mu = \frac{\alpha}{\beta}, \sigma = \frac{\sqrt{\alpha}}{\beta} \implies \frac{\sigma}{|\mu|} = \frac{1}{\sqrt{\alpha}}$$

Then, for a coefficient of variation of 2, we must have $\alpha = \frac{1}{4}$ in the prior distribution. Then, in order to get coefficient of variation 0.1, we must have $\alpha = 100$, and thus need at least $100 - \frac{1}{4}$ more samples, ie simply 100 more samples.

7.3.24

Each of the following will give the probability function and the corresponding a, b, c, d as the textbook, as in

$$f(x \mid \theta) = a(\theta)b(x)e^{c(\theta)d(x)}$$

a

$$f(x \mid p) = p^x(1-p)^{1-x} = (1-p)(\frac{p}{1-p})^x$$
$$a(p) = (1-p)$$
$$b(x) = 1$$
$$c(p) = \log(\frac{p}{1-p})$$
$$d(x) = x$$

b

$$f(x \mid \lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$$
$$a(\lambda) = e^{-\lambda}$$
$$b(x) = \frac{1}{x!}$$
$$c(\lambda) = \log(\lambda)$$
$$d(x) = x$$

c

$$f(x \mid p) = \binom{r+x-1}{x} p^r(1-p)^x$$
$$a(p) = p^r$$
$$b(x) = \binom{r+x-1}{x}$$
$$c(p) = \log(1-p)$$
$$d(x) = x$$

d

$$
\begin{aligned}
f(x \mid \mu) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\
a(\mu) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\mu^2}{2\sigma^2}} \\
b(x) &= e^{-\frac{x^2}{2\sigma^2}} \\
c(\mu) &= \frac{\mu}{2\sigma^2} \\
d(x) &= x
\end{aligned}
$$

e

$$
\begin{aligned}
f(x \mid \sigma^2) &= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\
a(\sigma^2) &= \frac{1}{\sqrt{2\pi\sigma^2}} \\
b(x) &= 1 \\
c(\sigma^2) &= -\frac{1}{2\sigma^2} \\
d(x) &= (x-\mu)^2
\end{aligned}
$$

f

$$
\begin{aligned}
f(x \mid \alpha) &= \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \\
a(\alpha) &= \frac{\beta^\alpha}{\Gamma(\alpha)} \\
b(x) &= e^{\beta x} \\
c(\alpha) &= \alpha - 1 \\
d(x) &= \log(x)
\end{aligned}
$$

g

$$
\begin{aligned}
f(x \mid \beta) &= \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \\
a(\beta) &= \frac{\beta^\alpha}{\Gamma(\alpha)} \\
b(x) &= x^{\alpha-1} \\
c(\beta) &= -\beta \\
d(x) &= x
\end{aligned}
$$

h

$$
\begin{aligned}
f(x \mid \alpha) &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \\
a(\alpha) &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \\
b(x) &= (1-x)^{\beta-1} \\
c(\alpha) &= \alpha - 1 \\
d(x) &= \log(x)
\end{aligned}
$$

i

$$
\begin{aligned}
f(x \mid \beta) &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \\
a(\beta) &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \\
b(x) &= x^{\alpha-1} \\
c(\beta) &= \beta - 1 \\
d(x) &= \log(1-x)
\end{aligned}
$$

7.4.3

a

We are minimizing the variance of the posterior distribution, which is beta with $\alpha = 5+y, \beta = 30-y$, which yields

$$
V = \frac{(5+y)(30-y)}{35^2(36)}
$$

This is maximized when $(5+y)(30-y)$ is maximized, which occurs at the axis of symmetry at $y = (30-5)/2 = 12.5 \implies y = 12, y = 13$ yields the same maximal mean squared error.

b

The variance is minimized when the numerator is minimized; this occurs at the endpoints of the parabola. Further, since $y = 0$ is further from the axis of symmetry, it must be minimized at $y = 0$.

7.4.10

This is a continuation of an earlier problem, and we know that the prior is gamma with $\alpha/\beta = 0.2, \alpha/\beta^2 = 1 \implies \beta = 0.2, \alpha = 0.04$, such that now, the posterior has $\alpha = 20.04, \beta = 20(3.8) + 0.2 = 76.2$. The mean of this distribution is $\alpha/\beta = 0.263$.

7.5.2

We have from the textbook that the MLE is $\overline{x_n} = 58/70$

7.5.7

We have the likelihood function

$$f(x \mid \beta) = \prod_{i=1}^{n} \beta e^{-\beta x_i} = \beta^n e^{-\beta \sum x_i}$$

Taking the logarithm,

$$\log(f(x \mid \beta)) = n\log(\beta) + -\beta \sum x_i$$

The maximizing condition is

$$\frac{n}{\beta} - \sum x_i = 0 \implies \beta = \frac{n}{\sum x_i} = \frac{1}{\overline{x_n}}$$

7.5.13

We have the following joint likelihood function

$$f(x, y \mid \mu_1, \mu_2) \propto e^{\sum_{i=1}^{n}[(\frac{x_i - \mu_x^2}{\sigma_i})^2 - 2\rho(\frac{x_i - \mu}{\sigma_x})(\frac{y_i - \mu_y}{\sigma_y}) + (\frac{y_i - \mu_y}{\sigma_y})^2]}$$

The first order conditions are then

$$\begin{aligned}\frac{\partial L}{\partial \mu_x} &\propto \frac{1}{\sigma_x^2}\left(\sum x_i - n\mu_x\right) - \frac{\rho}{\sigma_x\sigma_y}\left(\sum y_i - n\mu_y\right)\\ \frac{\partial L}{\partial \mu_y} &\propto \frac{1}{\sigma_y^2}\left(\sum y_i - n\mu_y\right) - \frac{\rho}{\sigma_x\sigma_y}\left(\sum x_i - n\mu_x\right)\end{aligned}$$

Note that we can pick the obvious choice $\mu_x = \overline{x}_n, \mu_y = \overline{y}_n$, which see both partials vanishing, suggesting that these are the MLE's.

7.6.2

The mean is equal to the variance in Poisson distributions, and the MLE of the mean (and therefore the variance) is $\sqrt{\overline{x}_n}$.

7.6.4

Suppose that the probability of a lamp failing in the T hours is p. Then, the likelihood function is

$$f(x \mid p) = p^x(1-p)^{n-x}$$

which was shown earlier to have MLE x/n. Since the above probability is exponentially distributed (that is, $p = 1 - e^{-\beta T}$), we can see that the MLE of β is

$$1 - e^{-\beta T} = \frac{x}{n} \implies \beta = -\frac{\log(1 - \frac{x}{n})}{T}$$

7.6.12

We wish to show here that $\hat{\beta} \xrightarrow{p} \beta$. We have that for $X_1, \ldots, X_n$, that the MLE β is $\hat{\beta} = \frac{1}{\overline{x}_n}$. The law of large numbers yields that $\lim_{n\to\infty} \overline{x}_n = \mu = \frac{1}{\beta}$. Then, $\hat{\beta} \xrightarrow{p} \frac{1}{\frac{1}{\beta}} = \beta$, which was what we wanted.

7.6.20

The method of moments suggests that $m_1 = \mu_1(\theta) = \theta$, which yields that the MLE and the method of moments gives the same result.

7.7.1

Not covered, skipped.

7.7.5

Not covered, skipped.

7.7.8

Not covered, skipped.

7.7.17

Not covered, skipped.

7.8.5

Not covered, skipped.

7.8.9

Not covered, skipped.

7.9.7

Not covered, skipped.

7.9.12

Not covered, skipped.