Quiz 1 solution
Math 5080-2
Sept. 2, 2015

1. Let $X_1$ and $X_2$ be independent $N(0,1)$ random variables. Each has density

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad -\infty < x < \infty.$$

Find the joint density $g(y_1, y_2)$ of $Y_1$ and $Y_2$, where

$$\begin{aligned} Y_1 &= X_1 + X_2, \\ Y_2 &= X_1 - X_2. \end{aligned} \tag{1}$$

Show all your work, and simplify your final result algebraically.

Hint: To find the inverse transformation, add and subtract the equations in (1).

Sol. The inverse transformation is

$$\begin{aligned} X_1 &= (Y_1 + Y_2)/2, \\ X_2 &= (Y_1 - Y_2)/2, \end{aligned}$$

which has Jacobian

$$J = \det \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix} = -1/2.$$

So the joint density of $Y_1$ and $Y_2$ is

$$\begin{aligned} g(y_1, y_2) &= f(x_1)f(x_2)|J| = f((y_1 + y_2)/2)f((y_1 - y_2)/2)|J| \\ &= \frac{1}{\sqrt{2\pi}} e^{-(y_1+y_2)^2/8} \frac{1}{\sqrt{2\pi}} e^{-(y_1-y_2)^2/8} \frac{1}{2} \\ &= \frac{1}{\sqrt{4\pi}} e^{-y_1^2/4} \frac{1}{\sqrt{4\pi}} e^{-y_2^2/4}. \end{aligned}$$

It is not necessary to mention it, but this shows that $Y_1$ and $Y_2$ are independent $N(0,2)$ random variables.

Quiz 2
Math 5080-2 Name: solutions
Sept. 9, 2015

1. Let $X_1, X_2, X_3$ be independent uniform$(0,1)$ random variables, and let $Y_1, Y_2, Y_3$ be their order statistics.

(a) Find the pdf of $Y_1 = \min(X_1, X_2, X_3)$.

Method 1. Use the formula derived in class:

$$f_{Y_1}(y_1) = \binom{3}{0,1,2} y_1^0 1 (1-y_1)^2 = 3(1-y_1)^2.$$

Method 2. Use the cdf method:

$$F_{Y_1}(y_1) = 1 - P(Y_1 > y_1) = 1 - P(X_1 > y_1)^3 = 1 - (1-y_1)^3,$$

so $f_{Y_1}(y_1) = 3(1-y_1)^2$.

Method 3. Integrate the joint pdf: $f(y_1, y_2, y_3) = 6,\ 0 < y_1 < y_2 < y_3 < 1$, so

$$f_{Y_1}(y_1) = \int_{y_1}^1 \int_{y_2}^1 6\, dy_3\, dy_2 = \int_{y_1}^1 6(1-y_2)\, dy_2 = 3(1-y_1)^2.$$

(b) Find the mean of $Y_1$. If you are not confident about your answer to (a), you may use the fact that

$$E[Y_1] = \int_0^1 P(Y_1 > y)\, dy.$$

$$E[Y_1] = \int_0^1 y_1\, 3(1-y_1)^2 dy_1 = 3\int_0^1 (y_1 - 2y_1^2 + y_1^3)\, dy_1 = 3\left(\frac{1}{2} - \frac{2}{3} + \frac{1}{4}\right) = \frac{1}{4}$$

or

$$E[Y_1] = \int_0^1 P(Y_1 > y)\, dy = \int_0^1 P(X_1 > y)^3 dy = \int_0^1 (1-y)^3 dy = \int_0^1 z^3 dz = \frac{1}{4}.$$

Quiz 3
Math 5080-2 Name:
Sept. 16, 2015

1. Consider a random sample of size $n$ from a distribution with CDF $F(x) = 1 - 1/(1+x)$ if $x > 0$, $= 0$ otherwise.

(a) Find the CDF $F_n(x)$ of $X_{1:n} = \min(X_1, X_2, \ldots, X_n)$.

Sol. $F_n(x) = 1 - (1 - F(x))^n = 1 - 1/(1+x)^n$ for $x > 0$.

(b) Use (a) to find the CDF $G_n(y)$ of $Y_n = n\, X_{1:n}$.

Sol. $G_n(y) = F_n(y/n) = 1 - 1/(1 + y/n)^n$ for $y > 0$.

(c) Use (b) to find the limiting CDF $G(y)$ of $Y_n = n\, X_{1:n}$ as $n \to \infty$.

Sol. $G(y) = 1 - 1/e^y = 1 - e^{-y}$ for $y > 0$. The limiting distribution is exponential.

Quiz 4
Math 5080-2 Name: solution
Sept. 23, 2015

1. Let $X_1, X_2, \ldots$ be independent exponential (density $f(x) = e^{-x}$, $x > 0$). We know that the sample median $X_{(n+1)/2\,:\,n}$ (if $n$ is odd) is asymptotically normal.

(a) Find the asymptotic mean and variance. Recall formulas $m = x_p$ (asymptotic mean is $p$th quantile) and $c^2/n$ (asymptotic variance) with $c^2 = p(1-p)/[f(x_p)]^2$; here $p = 1/2$.

Sol. Must find $x_p$. CDF is $F(x) = 1 - e^{-x}$, so solve $F(x_p) = p$. $1 - e^{-x_p} = p$ or $e^{-x_p} = 1-p$. $x_p = -\ln(1-p)$. With $p = 1/2$, we get $x_{1/2} = -\ln(1/2) = \ln 2$. Thus, asymptotic mean is $m = \ln 2$. And $c^2 = (1/4)/(e^{-\ln 2})^2 = (1/4)/(1/2)^2 = 1$. Thus, asymptotic variance is $1/n$.

(b) Explain why the reciprocal $1/X_{(n+1)/2\,:\,n}$ is also asymptotically normal, and find the asymptotic mean and variance.

Sol. If $Y_n$ is asymp$N(m, c^2/n)$, then $g(Y_n)$ is asymp$N(g(m), c^2|g'(m)|/n)$. We apply this with $g(y) = 1/y$, hence $g'(y) = -1/y^2$. Conclude that $1/X_{(n+1)/2\,:\,n}$ is asymptotically normal with asymptotic mean $1/\ln 2$ and asymptotic variance $1/[n(\ln 2)^2]$.

Quiz 5
Math 5080-2 Name: answers
Sept. 30, 2015

1. Let $X_1, X_2, \ldots, X_n$ be independent $N(\mu, \sigma^2)$. State the distribution of

(a) $\overline{X}_n := (X_1 + \cdots + X_n)/n \sim N(\mu, \sigma^2/n)$

(b) $\sum_{i=1}^n (X_i - \mu)^2/\sigma^2 \sim \chi^2(n)$

(c) $\sum_{i=1}^n (X_i - \overline{X}_n)^2/\sigma^2 \sim \chi^2(n-1)$

(d) $\sqrt{n}\,(\overline{X}_n - \mu)/\sqrt{\frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2} \sim t(n-1)$

(e) $n(\overline{X}_n - \mu)^2/\big[\frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X}_n)^2\big] \sim F(1, n-1)$

Quiz 6
Math 5080-2 Name: solution
Oct. 21, 2015

1. Suppose we have a random sample $X_1, X_2, \ldots, X_n$ from the lognormal distribution with density function

$$f(x;\theta) = (2\pi\sigma^2)^{-1/2}x^{-1}e^{-(\log x-\mu)^2/(2\sigma^2)}, \quad x > 0,$$

where $\theta = (\mu, \sigma^2)$. Find the maximum likelihood estimators of $\mu$ and $\sigma^2$.

Hints: Maximize the log likelihood by taking partial derivatives. First find $\hat{\mu}$, then find $\hat{\sigma}^2$. The argument is very similar to what we did in the $N(\mu, \sigma^2)$ case. (You may use log or ln, as you prefer.)

Sol.

$$L(\mu, \sigma^2) = (2\pi\sigma^2)^{-n/2}(x_1x_2\cdots x_n)^{-1}\exp\bigg\{-\frac{1}{2\sigma^2}\sum_{i=1}^n(\log x_i - \mu)^2\bigg\}$$

$$\log L(\mu, \sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \log(x_1x_2\cdots x_n) - \frac{1}{2\sigma^2}\sum_{i=1}^n(\log x_i - \mu)^2$$

$$\frac{\partial}{\partial\mu}\log L(\mu, \sigma^2) = \frac{1}{\sigma^2}\sum_{i=1}^n(\log x_i - \mu) = 0 \tag{1}$$

$$\frac{\partial}{\partial\sigma^2}\log L(\mu, \sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^n(\log x_i - \mu)^2 \tag{2}$$

Bu (1), $\hat{\mu} = \frac{1}{n}\sum_{i=1}^n \log x_i$. By (2), $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n(\log x_i - \hat{\mu})^2$.

Quiz 7
Math 5080-2 Name:
Oct. 28, 2015

1. Suppose we have a random sample $X_1, X_2, \ldots, X_n$ from the Poisson($\lambda$) distribution with density function

$$f(x;\lambda) = e^{-\lambda}\frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots,$$

where $\lambda > 0$. Recall that the mean and variance of a Poisson($\lambda$) random variable are both equal to $\lambda$.

(a) Find the Cramér-Rao lower bound (CRLB) for the variance of an unbiased estimator $T$ of $\tau(\lambda) = e^{-\lambda}$.

Sol. $\log f(x;\lambda) = -\lambda + x\log\lambda$, so $(\partial/\partial\lambda)\log f(x;\lambda) = -1 + x/\lambda$. Hence $I_1(\lambda) = E[(-1 + X/\lambda)^2] = \lambda^{-2}E[(X-\lambda)^2] = \lambda^{-2}\text{Var}(X) = \lambda^{-2}\lambda = \lambda^{-1}$. The CRLB is

$$(\tau'(\lambda))^2/(nI_1(\lambda)) = (-e^{-\lambda})^2/(n/\lambda) = \lambda e^{-2\lambda}/n.$$

(b) Notice that the estimator

$$T = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i=0\}}$$

(which is the proportion of observations that are equal to 0) is unbiased. Find its variance. Does it achieve the lower bound?

Sol. T is a sample proportion, so its variance is $p(1-p)/n$ with $p = P(X = 0) = e^{-\lambda}$. Thus, the variance $e^{-\lambda}(1-e^{-\lambda})/n$ does not achieve the CRLB.

Quiz 8
Math 5080-2 Name: Solution
Nov. 4, 2015

1. Let $X_1, \ldots, X_n$ be a random sample from EXP$(1/\theta)$, which has density $f(x;\theta) = \theta e^{-\theta x}$, $x > 0$. Now assume that $\Theta$ is random with prior density EXP(1), that is, $f_\Theta(\theta) = e^{-\theta}$, $\theta > 0$.

(a) Find the posterior density of $\Theta$, given $x_1, \ldots, x_n$, up to a constant (in $\theta$) multiple. Is this density of the form GAM$(\alpha, \beta)$, which has the form (as a function of $\theta$)

$$f(\theta) = \text{constant } \theta^{\beta-1} e^{-\theta/\alpha} \tag{1}$$

If so, find the parameters $\alpha$ and $\beta$ in terms of $x_1, \ldots, x_n$ and $n$.

Sol. The posterior density is proportional to

$$e^{-\theta} \prod_{i=1}^{n} \theta e^{-\theta x_i} = \theta^n e^{-\theta(1+\sum x_i)}$$

This has the form of (1) with $\beta - 1 = n$ and $1/\alpha = 1 + \sum x_i$, or $\alpha = 1/(1 + \sum x_i)$ and $\beta = 1 + n$.

(b) Find the Bayes estimator of $\Theta$, assuming squared error loss. Hint: The mean of the GAM$(\alpha, \beta)$ density in (1) is the product $\alpha\beta$.

Sol. The Bayes estimator is the mean of the posterior distribution, which is $\alpha\beta$ or

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

Quiz 9
Math 5080-2 Name:
Nov. 11, 2015

1. Let $X_1, X_2, \ldots, X_n$ be a random sample from the two-parameter exponential distribution, with density

$$f(x;\theta,\eta) = \frac{1}{\theta} e^{-(x-\eta)/\theta} I_{(\eta,\infty)}(x),$$

where $I_{(\eta,\infty)}(x) = 1$ if $x > \eta$ and $= 0$ otherwise. Using the factorization theorem, find a pair of jointly sufficient statistics for $(\theta, \eta)$.

Sol.

$$\prod_{i=1}^{n} f(x_i;\theta,\eta) = \frac{1}{\theta^n} e^{-\sum_{i=1}^{n}(x_i-\eta)/\theta} \prod_{i=1}^{n} I_{(\eta,\infty)}(x_i) = \frac{1}{\theta^n} e^{-(\sum_{i=1}^{n} x_i - n\eta)/\theta} I_{(\eta,\infty)}(x_{1:n})$$

We find that $\sum_{i=1}^{n} x_i$ and $x_{1:n}$ are jointly sufficient. Equivalently, $\overline{x}$ and $x_{1:n}$ are jointly sufficient.

Quiz 10
Math 5080-2 Name: Sol.
Dec. 2, 2015

1. Let $X_1, X_2, \ldots, X_n$ be a random sample from a WEI$(\theta, 2)$ distribution with density

$$f(x;\theta) = \frac{2}{\theta^2} x e^{-x^2/\theta^2}, \quad x > 0.$$

(a) Find a pivotal quantity, that is, an expression depending on the sample and the parameter $\theta$ whose distribution does not depend on $\theta$. Hint 1: Is $\theta$ a scale parameter? Hint 2: The maximum likelihood estimator of $\theta$ is

$$\hat{\theta} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}.$$

Sol. Yes, $\theta$ is a scale parameter, since $f(x;\theta) = \frac{1}{\theta} 2 \frac{x}{\theta} e^{-(x/\theta)^2}$. So $\hat{\theta}/\theta$ is a pivotal quantity if $\hat{\theta}$ is the MLE. We find that

$$\frac{\hat{\theta}}{\theta} = \frac{1}{\theta}\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}.$$

is pivotal.

(b) Use this to find a $100(1-\alpha)\%$ confidence interval for $\theta$. (You may denote the $\gamma$-quantile of the pivotal quantity by $q_\gamma$. You need not evaluate it.)

Sol.

$$P\left(q_{\alpha/2} < \frac{\hat{\theta}}{\theta} < q_{1-\alpha/2}\right) = 1 - \alpha,$$

so

$$P\left(\frac{\hat{\theta}}{q_{1-\alpha/2}} < \theta < \frac{\hat{\theta}}{q_{\alpha/2}}\right) = 1 - \alpha.$$

The interval

$$\left(\frac{\hat{\theta}}{q_{1-\alpha/2}}, \frac{\hat{\theta}}{q_{\alpha/2}}\right)$$

is a $100(1-\alpha)\%$ confidence interval for $\theta$.

Quiz 11
Math 5080-2 Name:
Dec. 9, 2015

1. We have two independent normal samples of sizes $n_1 = 3$ and $n_2 = 4$, from $N(\mu_1, \sigma^2)$ and $N(\mu_2, \sigma^2)$ (note the equal but unknown variances). The data for the $x$-sample are $1, 3, 5$, and the data for the $y$ sample are $2, 4, 6, 8$. (Admittedly, this is a bit artificial, but numbers were chosen so that you can do computations by hand.) Write down a 90% confidence interval for the mean difference $\mu_2 - \mu_1$. You may use your book (or borrow one) to look up the relevant quantile in a table.

Sol. The general form of the $100(1-\alpha)\%$ confidence interval was derived as

$$\overline{y} - \overline{x} \pm t_{1-\alpha/2}(n_1 + n_2 - 2) s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

with $s_p^2 = [(n_1 - 1)s_x^2 + (n_2 - 1)s_y^2]/(n_1 + n_2 - 2)$.

We can compute each of these quantities. $\overline{x} = 3$ and $\overline{y} = 5$, hence $(n_1 - 1)s_x^2 = 8$ and $(n_2 - 1)s_y^2 = 20$. We get $s_p^2 = 28/5 = 5.6$. We look up $t_{0.95}(5)$, getting 2.015. Result is

$$2 \pm 2.015\sqrt{5.6}\sqrt{\frac{1}{3} + \frac{1}{4}} = 2 \pm 3.6419.$$

or $(-1.6419, 5.6419)$.