1. MGB III.2

(a) Find the mode of the beta distribution

Solution:

The density function is given by

$$f(x|a,b) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} I_{(0,1)}(x),$$

whose derivative is

$$\begin{aligned} f'(x|a,b) &= \frac{1}{B(a,b)} \Big[(a-1)x^{a-2}(1-x)^{b-1} - (b-1)x^{a-1}(1-x)^{b-2}\Big] I_{(0,1)}(x) \\ &= \frac{1}{B(a,b)} x^{a-2}(1-x)^{b-2}\big[(a-1)(1-x) - (b-1)x\big] I_{(0,1)}(x) \\ &= \frac{1}{B(a,b)} x^{a-2}(1-x)^{b-2}\big[x(2-a-b) + (a-1)\big] I_{(0,1)}(x), \end{aligned}$$

so long as $a \neq 1$ and $b \neq 1$ in which case $f(x) = I_{(0,1)}$ and the mode is clearly 1. Otherwise, $x = 0, 1$ are critical points, as well as $(a-1)/(a+b-2)$ when both $a < 1$ and $b < 1$ or when $a > 1$ and $b > 1$. When $a < 1$ or when $b < 1$, then both x^{1-a} and x^{1-b} are unbounded and, thus, the mode does not exist. When both $a, b > 1$ then $x(2-a-b) + (a-1)$ is a line with negative slope $(2-a-b)$, hence at the critical point $(a-1)/(a+b-2)$ f attains a maximum on $(0,1)$ by the first derivative test.

□

(b) Find the mode of the gamma distribution

Solution The density is given by

$$f(x|r,\lambda) = \frac{\lambda}{\Gamma(r)}(\lambda x)^{r-1} e^{-\lambda x} I_{(0,\infty)}(x) \quad \text{for } \lambda, r > 0,$$

whose derivative is given by

$$\begin{aligned} f'(x|r,\lambda) &= \frac{\lambda}{\Gamma(r)}\Big[(r-1)\lambda(\lambda x)^{r-2}e^{-\lambda x} - \lambda(\lambda x)^{r-1}e^{-\lambda x}\Big] \\ &= \frac{\lambda^2}{\Gamma(r)}(\lambda x)^{r-2}e^{-\lambda x}\Big[(r-1) - \lambda x\Big]. \end{aligned}$$

Note that when $r < 1$, f is unbounded and thus no mode exists. When $r = 1$, $f'(x) < 0$ for all $x \in (0, \infty)$ so f is maximized as $x \to 0$, hence there is no mode since $0 \notin (0, 1)$. Otherwise, $x = \frac{r-1}{\lambda}$ is a critical point. Moreover, $(r-1) - \lambda x$ is a line with negative slope $-\lambda$ and by the first derivative test, f attains its maximum at $\frac{r-1}{\lambda}$ and, thus, is a mode.

□

2. MGB III.15 Let X be normally distributed with mean μ and variance σ^2. Truncate the density of X on the left at a and the right at b, and then calculate the mean of the truncated distribution. (Note that the mean of the truncated distribution should fall between a and b. Furthermore, if $a = \mu - c$ and $b = \mu + c$, then the mean of the truncated distribution is μ.

Solution:

The truncation, Y, is distributed with the p.d.f.

$$\frac{\phi_{\mu,\sigma^2}(y)I_{(a,b)}(y)}{\Phi_{\mu,\sigma^2}(b) - \Phi_{\mu,\sigma^2}(a)}.$$

So the mean is

$$\begin{aligned}
E(Y) &= \frac{1}{\Phi_{\mu,\sigma^2}(b) - \Phi_{\mu,\sigma^2}(a)} \int_a^b y\phi_{\mu,\sigma^2}(y)dy \\
&= \frac{1}{\Phi_{\mu,\sigma^2}(b) - \Phi_{\mu,\sigma^2}(a)} \int_a^b \frac{y}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right) dy.
\end{aligned}$$

Let $u = (y - \mu)/\sigma$ then $y/\sigma = u - \mu/\sigma$ and $dy = \sigma du$, so continuing from above,

$$\begin{aligned}
E(Y) &= \frac{1}{\Phi_{\mu,\sigma^2}(b) - \Phi_{\mu,\sigma^2}(a)} \int_{(a-\mu)/\sigma}^{(b-\mu)/\sigma} \Big(\sigma u\phi(u) - \mu\phi(u)\Big)du \\
&= \frac{1}{\Phi_{\mu,\sigma^2}(b) - \Phi_{\mu,\sigma^2}(a)} \int_{(a-\mu)/\sigma}^{(b-\mu)/\sigma} \sigma u\phi(u)du - \mu\left[\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)\right] \\
&= \frac{\sigma}{\sqrt{2\pi}(\Phi_{\mu,\sigma^2}(b) - \Phi_{\mu,\sigma^2}(a))} \int_{(a-\mu)/\sigma}^{(b-\mu)/\sigma} ue^{-u^2/2}du + \mu.
\end{aligned}$$

Let $w = u^2/2$, then $dw = udu$ and continuing from above

$$\begin{aligned}
E(Y) &= \frac{\sigma^2}{\Phi_{\mu,\sigma^2}(b) - \Phi_{\mu,\sigma^2}(a)} \cdot \frac{1}{\sigma\sqrt{2\pi}} \int_{(a-\mu)^2/(2\sigma^2)}^{(b-\mu)^2/(2\sigma^2)} e^{-w}dw + \mu. \\
&= \sigma^2 \frac{\phi_{\mu,\sigma}(a) - \phi_{\mu,\sigma}(b)}{\Phi_{\mu,\sigma^2}(b) - \Phi_{\mu,\sigma^2}(a)} + \mu.
\end{aligned}$$

□

3. MGB III.17 Let X be the life in hours of a radio tube. Assume that X is normally distributed with mean 200 and variance σ^2. If a purchaser of such radio tubes requires that at least 90 percent of the tubes have lives exceeding 150 hours, what is the largest value of σ can be and still have the purchaser satisfied?

Solution:

Since $X \sim N(200, \sigma^2)$, we have that

$$\begin{aligned} P(X \geq 150) &= 1 - \Phi_{200,\sigma^2}(150) \\ &= 1 - \Phi\left(\frac{150-200}{\sigma^2}\right) \\ &= \Phi\left(\frac{50}{\sigma^2}\right) \geq .9 \end{aligned}$$

Note that Φ is strictly increasing and bounded between 0,1, hence it has a unique value z^* such that $\Phi(z^*) = .9$ and $\Phi(z) \geq .9$ for all $z \geq z^*$. Thus, we require $\frac{50}{\sigma^2} \geq z*$ or equivalently $\sigma^2 \leq \frac{50}{z^*}$ since both z^* and σ^2 are greater than 0. We can obtain a numerical estimate for the upper bound on the variance with the program `R` using the command `50/qnorm(.9)` which yields `39.01521`.

□

4. MGB III.19a The distribution given by

$$f(x|\beta) = \frac{1}{\beta^2} x e^{-\frac{1}{2}(x/\beta)^2} I_{(0,\infty)}(x) \quad \text{for } \beta > 0$$

is called the *Raleigh* distribution. Show that the mean and variance exist and find them.

Solution:

The nth moment of X is given by

$$\begin{aligned} E(X^n) &= \int_0^\infty \frac{1}{\beta^2} x^{n+1} e^{-\frac{1}{2}(x/\beta)^2}\, dx \\ &= \int_0^\infty \beta^n (2u)^{\frac{n}{2}} e^{-u} du \qquad \text{where } u = \frac{1}{2}(x/\beta)^2 \text{ and } du = x/\beta^2 dx \\ &= \beta^n 2^{\frac{n}{2}} \Gamma\left(1 + \frac{n}{2}\right). \end{aligned}$$

Thus $\mu = E(X) = \frac{2\beta}{\sqrt{2}}\sqrt{\pi} = \beta\sqrt{\frac{\pi}{2}}$ and $\sigma^2 = E(X^2) - E(X)^2 = 4\beta^2 - \beta^2\frac{\pi}{2} = \beta^2\frac{4-\pi}{2}$.

□

5. MGB III.20a The distribution given by

$$F(x|\beta) = \frac{4}{\beta^3\sqrt{\pi}} x^2 e^{-x^2/\beta^2} I_{(0,\infty)}(x) \quad \text{for } \beta > 0$$

is called the *Maxwell* distribution. Show that the mean and variance exist and find them.

Solution:

The nth moment of X is given by

$$\begin{aligned} E(X^n) &= \int_0^\infty \frac{4}{\beta^3\sqrt{\pi}} x^{n+2} e^{-(x/\beta)^2}\, dx \\ &= \frac{2}{\sqrt{\pi}} \int_0^\infty \beta^n (u)^{\frac{n+1}{2}} e^{-u} du \qquad \text{where } u = (x/\beta)^2 \text{ and } du = 2x/\beta^2 dx \\ &= \frac{2\beta^n}{\sqrt{\pi}} \Gamma\left(\frac{n+3}{2}\right). \end{aligned}$$

Thus $\mu = E(X) = \frac{2\beta}{\sqrt{\pi}}$ and $\sigma^2 = E(X^2) - E(X)^2 = 2\beta^2\Gamma(5/2) - \frac{4\beta^2}{\pi} = \beta^2(3/4 - 4/\pi)$.

□

6. MGB III.28 Show that

$$P(X \geq k) = \sum_{x=k}^{n} \binom{n}{x} p^x (1-p)^{n-x} = \frac{1}{B(k, n-k+1)} \int_0^p u^{k-1}(1-u)^{n-k} du$$

for X a binomially distributed random variable. That is, if X is binomially distributed with parameters n and p and Y is beta-distributed with parameters k and $n-k+1$, then $F_Y(p) = 1 - F_X(k-1)$.

Solution Let us evaluate the integral on the far right hand side via integration by parts $n-k$ times. That is,

$$\begin{aligned} &\frac{1}{B(k, n-k+1)} \int_0^p u^{k-1}(1-u)^{n-k} du \\ &= \frac{1}{B(k, n-k+1)} \left((1-u)^{n-k} \frac{u^k}{k}\Big|_0^p + \int_0^p \frac{u^k}{k}(n-k)(1-u)^{n-(k+1)}\, du \right) \\ &\cdots \\ &= \frac{1}{B(k, n-k+1)} \sum_{x=k}^{n} \left[\frac{(n-k)!}{(n-x)!}(1-p)^{n-k} \right] \cdot \left[\frac{(k-1)!}{x!} p^k \right] \\ &= \frac{\Gamma(n+1)}{\Gamma(k)\Gamma(n-k+1)} \sum_{x=k}^{n} \left[\frac{(n-k)!}{(n-x)!}(1-p)^{n-k} \right] \cdot \left[\frac{(k-1)!}{x!} p^k \right] \\ &= \frac{n!}{(k-1)!(n-k)!} \sum_{x=k}^{n} \left[\frac{(n-k)!}{(n-x)!}(1-p)^{n-k} \right] \cdot \left[\frac{(k-1)!}{x!} p^k \right] \\ &= \sum_{x=k}^{n} \binom{n}{x} p^x (1-p)^{n-x} = P(X \geq k) \quad \square \end{aligned}$$

7. MGB V.10 A certain explosive device will detonate if any one of n short-lived fuses lasts longer than .8 seconds. Let X_i represent the life of the ith fuse. It can be assumed that each X_i is uniformly distributed over the interval 0 to 1 second. Furthermore, it can be assumed that the X_i's are independent.

(a) How many fuses are needed (i.e. how large should n be) if one wants to be 95 percent certain that the device will detonate?

Solution:

The device will detonate if $\max_{i=1..n}(X_i) \geq .8$ so we consider

$$\begin{aligned} P\left(\max_{i=1..n}(X_i) \geq .8\right) &= 1 - P\left(\max_{i=1..10} X_i < .8\right) \\ &= 1 - \prod_{i}^{10} P(X_i < .8) && \text{by independence} \\ &= 1 - (.8)^n. \end{aligned}$$

If we require that this event has at least a probability of .95 then $1 - (.8)^n \geq .95$ if and only if $n \log .8 \leq \log .05$ if and only if $n \geq \frac{\log .05}{\log .8} \approx 13.43$. Hence, at least 14 fuses guarantees a probability of .95 of detonation.

□

(b) If the device has nine fuses what is the average life of the fuse that lasts the longest?

Solution:

The c.d.f. of the random variable $\max_{i=1..n}(X_i)$ is given by $F(x) = P\left(\max_{i=1..n} X_i < x\right) = x^n$. So, the p.d.f. is $F'(x) = nx^{n-1}$, and the mean is given by the integral

$$\int_0^1 nx^n = \frac{n}{n+1}.$$

Hence, in the case of nine fuses the mean is 9/10.

□

8. MGB V.13 Let X_1 and X_2 be independent standard normal random variables. Let $Y_1 = X_1 + X_2$ and $Y_2 = X_1^2 + X_2^2$.

(a) Show that the join moment generating function of Y_1 and Y_2 is

$$\frac{\exp\left[t_1^2/(1-2t_2)\right]}{1-2t_2} \quad \text{for } -\infty < t_1 < \infty \text{ and } -\infty < t_2 < \frac{1}{2}.$$

Solution:

The joint moment generating function is given by

$$\begin{aligned}
M_{X_1,X_2}(t_1,t_2) &= \iint_{\mathbb{R}} e^{t_1(x_1+x_2)+t_2(x_1^2+x_2^2)} \phi(x_1,x_2)\, dx_1 dx_2 \\
&= \iint_{\mathbb{R}} e^{t_1(x_1+x_2)} \frac{1}{2\pi} e^{(t_2-1/2)(x_1^2+x_2^2)}\, dx_1 dx_2 \\
&= \left(\sigma_{t_2} \int_{\mathbb{R}} e^{t_1 x_1} \frac{1}{\sqrt{2\pi}\sigma_{t_2}} e^{-(x_1/\sigma_{t_2})^2}\, dx_1\right) \left(\sigma_{t_2} \int_{\mathbb{R}} e^{t_1 x_2} \frac{1}{\sqrt{2\pi}\sigma_{t_2}} e^{-(x_2/\sigma_{t_2})^2}\, dx_2\right),
\end{aligned}$$

where $\sigma_{t_2}^2 = (2t_2-1)^{-1}$, and thus $\infty < t_2 < 1/2$. Note that each integral above the moment generating functions for the random variable $Y \sim N(0,\sigma^2)$. That is,

$$\begin{aligned}
M_{X_1,X_2}(t_1,t_2) &= (\sigma_{t_2} M_Y(t_1))^2 \\
&= \frac{\exp\left[t_1^2/(1-2t_2)\right]}{1-2t_2}.
\end{aligned}$$

□

(b) Find the correlation coefficient of Y_1 and Y_2.

Solution:

Recall the fact that

$$E\left(X_1^n X_2^m\right) = \lim_{(t_1,t_2)\to(0,0)} \frac{\partial^n \partial^m}{\partial t_1^n \partial t_2^m} M_{X_1,X_2}(t_1,t_2).$$

And note

$$\lim_{t_1\to 0} \frac{\partial}{\partial t_1} M_{X_1,X_2}(t_1,t_2) = \lim_{t_1\to 0} \frac{2t_1 e^{t_1^2/(1-2t_2)}}{(1-2t_2)^2} = 0.$$

Moreover, the partial derivative of M_{X_1,X_2} and $\partial/\partial t_1 M_{X_1,X_2}$ with respect to t_2 both involve the quotient rule so that the denominator is some power of $(2-t_2)$. Hence t_2 is not a singularity of those equations and thus $E(X_1)E(X_2) = 0$ and $E(X_1X_2) = 0$. Thus the correlation is $E(X_1X_2) - E(X_1)E(X_2) = 0$.

□

9. MGB V.22 Kitty Oil Co. has has decided to drill for oil in 10 different locations; the cost of drilling at each location is \$10,000. (Total cost is then \$100,000.) The probability of finding oil in a given location is only $\frac{1}{5}$, but if oil is found at a given location then the amount of money the company will get selling oil (excluding the initial \$10,000 drilling cost) from that location is an exponential random variable with mean \$50,000. Let Y be the random variable that denotes the number of locations where oil is found, and let Z denote the total amount of money received from selling oil from all the locations.

(a) Find $E(Z)$.

Solution:

Note that $Y \sim \text{Binomial}(10, \frac{1}{5})$ and that if each site $Z_i \sim \text{Exponential}\left(\frac{1}{50000}\right)$, then

$$(Z|Y=y) = \sum_{i=1}^{y} Z_i = \text{Gamma}\left(y, \frac{1}{50000}\right). \quad \text{(Thanks to a hint from Solomon.)}$$

We can use this expression and Adam's Law to calculate the expectation of Z. That is, if $\lambda = \frac{1}{50000}$ then

$$\begin{aligned} E(Z) &= E_Y\left(E\left(Z|Y\right)\right) \\ &= E_Y\left(\frac{Y}{\lambda}\right) \\ &= \frac{10 \cdot \frac{1}{5}}{50000} = 10000. \end{aligned}$$

□

(b) Find $P(Z > 100,000|Y = 1)$ and $P(Z > 100,000|Y = 2)$.

Solution:

From (a), $(Z|Y=1) \sim \text{Gamma}\left(1, \frac{1}{50000}\right)$, so

$$\begin{aligned} P(Z > 100,000|Y=1) &= 1 - \left(1 - e^{-100,000\lambda}\right) \\ &= e^{-2} \approx 0.1353353. \end{aligned}$$

In general $P(Z > 100,000|Y = n) = 1 - F_X(100,000)$ where $X \sim \text{Gamma}(n, \frac{1}{50000})$. We can use the command `pgamma` in computer program `R` to obtain

$$(Z > 100,000|Y = n) = 1 - F_X(100,000) \approx 0.4060058.$$

□

(c) How would you find $P(Z > 100,000)$? Is $P(Z > 100,000) > \frac{1}{2}$.

Solution:

By the law of total probability and the definition of conditional probability,

$$P(Z > 100,000) = \bigcup_{y=1}^{10} P(Z > 100,000, Y = y) \quad = \bigcup_{y=1}^{10} P(Z > 100,000|Y = y)P(Y = y).$$

Continued on next page.

We can compute this with the following `R` code:

```
> sum((1-pgamma(100000,1:10,1/50000))*dbinom(1:10,10,1/5))
[1] 0.4019745
```

$\square$

10. MGB V.54 Let X_1 and X_2 be independent random variables, each normally distributed with parameters $\mu = 0$ and $\sigma^2 = 1$. Find the joint distribution of $Y_1 = X_1^2 + X_2^2$ and $Y_2 = X_1/X_2$. Find the marginal distribution of Y_1 and of Y_2. Are Y_1 and Y_2 independent?

Solution:

Let us decompose $\mathbb{R} \times \mathbb{R} = A_1 \cup A_2 \cup K$ where $A_1 = \big(\mathbb{R} \times (-\infty, 0)\big)$, $A_2 = \big(\mathbb{R} \times (0, \infty)\big)$, and $K = \{(x_1, x_2) | x_1 = 0 \text{ or } x_2 = 0\}$. Note that this union is disjoint and that the measure of K is zero. Moreover, if $g(x_1, x_2) = (x_1^2 + x_2^2, x_1/x_2)$, then g is invertible on A_1 by

$$g_1^{-1}(x_1, x_2) = \sqrt{\frac{y_1}{y_2^2 + 1}}(y_2, 1)$$

and on A_2 by

$$g_2^{-1}(x_1, x_2) = \sqrt{\frac{y_1}{y_2^2 + 1}}(-y_2, -1)$$

The is key step in verifying this is noting that $y_2 = \sqrt{y_2^2}$ when $y_2 > 0$ and $y_2 = -\sqrt{y_2^2}$ when $y_2 < 0$. Calculating each Jacobian is straight-forward, but lengthy. Summarized,

$$\left| \begin{pmatrix} \frac{\pm y_2}{2\sqrt{y_1}\sqrt{{y_2}^2+1}} & \frac{\pm\sqrt{y_1}\sqrt{{y_2}^2+1}}{{y_2}^4+2\,{y_2}^2+1} \\ \frac{1}{2\sqrt{y_1}\sqrt{{y_2}^2+1}} & -\frac{\sqrt{y_1}\,y_2}{({y_2}^2+1)^{\frac{3}{2}}} \end{pmatrix} \right| = \frac{1}{2y_2^2 + 2}.$$

Hence the joint distribution of (Y_1, Y_2) is

$$f_{Y_1,Y_2}(y_1, y_2) = \frac{1}{2y_2^2 + 2}\Big[f_{X_1,X_2}(g_1^{-1}(y_1, y_2)) + f_{X_1,X_2}(g_2^{-1}(y_1, y_2))\Big] I_{(0,\infty)\times\mathbb{R}}(y_1, y_2)$$

$$= \frac{1}{y_2^2 + 1}\left[\phi\left(\sqrt{\frac{y_1}{y_2^2 + 1}}y_2\right)\phi\left(\sqrt{\frac{y_1}{y_2^2 + 1}}\right)\right] I_{(0,\infty)\times\mathbb{R}}(y_1, y_2) \qquad \text{since } \phi \text{ is even,}$$

$$= \frac{1}{y_2^2 + 1}\left[\frac{1}{2\pi}\exp\left(-\left(\frac{y_1}{y_2^2 + 1}y_2^2 + \frac{y_1}{y_2^2 + 1}\right)/2\right)\right] I_{(0,\infty)\times\mathbb{R}}(y_1, y_2)$$

$$= \frac{1}{y_2^2 + 1}\left[\frac{1}{2\pi}e^{-y_1/2}\right] I_{(0,\infty)\times\mathbb{R}}(y_1, y_2).$$

Note that the p.d.f. factors, and thus Y_1 is independent of Y_2.