MATH 556 - EXERCISES 3: SOLUTIONS

1 We use the usual marginalization formula if X is continuous

$$f_Y(y) = \int f_{X,Y}(x,y)\,dx = \int f_{Y|X}(y|x) f_X(x)\,dx.$$

In each case, the resulting integrand is proportional to a pdf, so that the integral can be evaluated directly.

(i) For $y = 0, 1, 2, \ldots$,

$$\begin{aligned} f_Y(y) = \int_0^\infty f_{Y|X}(y|x) f_X(x)\,dx &= \int_0^\infty \frac{e^{-x}x^y}{y!}\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\,dx \\ &= \frac{\beta^\alpha}{y!\Gamma(\alpha)} \int_0^\infty x^{y+\alpha-1} e^{-(\beta+1)x}\,dx \\ &= \frac{\beta^\alpha}{y!\Gamma(\alpha)} \frac{\Gamma(y+\alpha)}{(\beta+1)^{y+\alpha}} = \frac{\Gamma(y+\alpha)}{\Gamma(y+1)\Gamma(\alpha)} \theta^y (1-\theta)^\alpha \end{aligned}$$

where $\theta = 1/(\beta+1)$. Hence Y has a negative-binomial type distribution, with parameters $\alpha > 0$ and θ where $0 < \theta < 1$.

(ii) For $y > 0$,

$$\begin{aligned} f_Y(y) = \int_0^\infty f_{Y|X}(y|x) f_X(x)\,dx &= \int_0^\infty x e^{-xy} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\,dx \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha+1-1} e^{-(y+\beta)x}\,dx \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{(y+\beta)^{\alpha+1}} = \frac{(\alpha+1)\beta^\alpha}{(y+\beta)^{\alpha+1}} \end{aligned}$$

Hence Y has a Pareto type distribution (see formula sheet), with parameters $\alpha, \beta > 0$.

(iii) For $0 \le y \le n$,

$$\begin{aligned} f_Y(y) = \int_0^1 f_{Y|X}(y|x) f_X(x)\,dx &= \int_0^1 \binom{n}{y} x^y (1-x)^{n-y} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}\,dx \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{y+\alpha-1}(1-x)^{n-y+\beta-1}\,dx \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(y+\alpha)\Gamma(n-y+\beta)}{\Gamma(n+\alpha+\beta)} \end{aligned}$$

2 In the mixture model, then in the continuous case,

$$\begin{aligned} E_{f_Y}[Y] &= \int y f_Y(y)\,dy = \int y \left\{ \sum_{k=1}^K \pi_k f_k(y|\theta_k) \right\} dy \\ &= \sum_{k=1}^K \pi_k \left\{ \int y f_k(y|\theta_k)\,dy \right\} = \sum_{k=1}^K \pi_k \mu_k \end{aligned}$$

as required. An identical calculation reveals that

$$M_Y(t) = \sum_{k=1}^K \pi_k M_k(t;\theta_k)$$

where $M_k(t;\theta_k)$ is the mgf for $f_k(y|\theta_k)$, for $k = 1, \ldots, K$.

3 The conditional distribution of X is discrete on $\{0,1\}$. By the definition of conditional probability (or by Bayes Theorem)

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \propto f_{Y|X}(y|x)f_X(x)$$

Now for $x = 0$,

$$f_{Y|X}(y|0) = f_1(y) = I_{\{0\}}(y) \qquad f_X(0) = \pi_1$$

and for $x = 1$,

$$f_{Y|X}(y|1) = f_2(y) = \frac{e^{-y^2/2}}{\sqrt{2\pi}} \qquad f_X(1) = 1 - \pi_1$$

say, using π_1 to distinguish this probability from the constant π. Then for $x = 0$

$$f_{X|Y}(0|0) = \frac{\pi_1}{\pi_1 + (1-\pi_1)e^{-y^2/2}/\sqrt{2\pi}}$$

and for $x = 1$

$$f_{X|Y}(1|0) = \frac{(1-\pi_1)e^{-y^2/2}/\sqrt{2\pi}}{\pi_1 + (1-\pi_1)e^{-y^2/2}/\sqrt{2\pi}} \quad .$$

4 With the correct skewness formula, we need to compute

$$\varsigma = \frac{E_{f_X}[(X-\mu)^3]}{\sigma^3} \qquad \kappa = \frac{E_{f_X}[(X-\mu)^4]}{\sigma^4}$$

and hence need the first four central moments. If the mgf, M_X, exists, then it yields expressions for the first four moments after differentiation.

(i) For the standard Normal

$$\begin{aligned}
M_X(t) &= e^{t^2/2} \\
M_X^{(1)}(t) &= te^{t^2/2} \qquad \therefore \qquad M_X^{(1)}(0) = 0 \\
M_X^{(2)}(t) &= (t^2+1)e^{t^2/2} \qquad \therefore \qquad M_X^{(2)}(0) = 1 \\
M_X^{(3)}(t) &= (t^3+3t)e^{t^2/2} \qquad \therefore \qquad M_X^{(3)}(0) = 0 \\
M_X^{(4)}(t) &= (t^4+6t^2+3)e^{t^2/2} \qquad \therefore \qquad M_X^{(4)}(0) = 3
\end{aligned}$$

so $\mu = 0$ and $\sigma^2 = 1$, and hence

$$\begin{aligned}
\varsigma &= \frac{E_{f_X}[(X-\mu)^3]}{\sigma^3} = \frac{E_{f_X}[X^3]}{\sigma^3} = 0 \\
\kappa &= \frac{E_{f_X}[(X-\mu)^4]}{\sigma^4} = \frac{E_{f_X}[X^4]}{\sigma^4} = 3.
\end{aligned}$$

Note: some texts define the kurtosis for a general pmf/pdf f_X as

$$\frac{E_{f_X}[(X-\mu)^4]}{\sigma^4} - 3, \tag{1}$$

*as $\kappa = 3$ for the standard Normal distribution. I will refer to the quantity in equation (1) as the **excess kurtosis**.*

(ii) Throughout, we assume that all expectations used are finite. Note that, given $V = v$, $X = \sqrt{v}Z$ where $Z \sim N(0,1)$ such that Z and V are independent. Hence, using iterated expectation, we have

$$\begin{aligned} E_{f_X}[X^r] &= E_{f_V}[E_{f_{X|V}}[X^r|V=v]] \\ &= E_{f_V}[E_{f_{Z|V}}[v^{r/2}Z^r|V=v]] \\ &= E_{f_V}[E_{f_Z}[Z^r]\, V^{r/2}] \\ &= E_{f_V}[V^{r/2}]E_{f_Z}[Z^r] \end{aligned}$$

the last line following by independence. Hence

$$\begin{aligned} r=1 &\quad : \quad E_{f_X}[X] = E_{f_V}[V^{1/2}]E_{f_Z}[Z] = 0 \\ r=2 &\quad : \quad E_{f_X}[X^2] = E_{f_V}[V]E_{f_Z}[Z] = E_{f_V}[V] \\ r=3 &\quad : \quad E_{f_X}[X^3] = E_{f_V}[V^{3/2}]E_{f_Z}[Z^3] = 0 \\ r=4 &\quad : \quad E_{f_X}[X^4] = E_{f_V}[V^2]E_{f_Z}[Z^4] = 3E_{f_V}[V^2] \end{aligned}$$

using the moments of a standard normal rv from above. Hence, as $\mu = 0$, and

$$\sigma^2 = E_{f_X}[X^2] = E_{f_V}[V]$$

we have

$$\varsigma = \frac{E_{f_X}[X^3]}{\sigma^3} = 0 \qquad\qquad \kappa = \frac{E_{f_X}[X^4]}{\sigma^4} = \frac{3E_{f_V}[V^2]}{\{E_{f_V}[V]\}^2}$$

Note that, unless V has a distribution with variance zero (which occurs if V has a degenerate distribution, $P[V = c] = 1$ for some c), $\kappa > 3$ (by Jensen's Inequality), so the scale mixture has a kurtosis value *greater* than that of the Normal, that is, the distribution is *leptokurtic* compared to the Normal.

(iii) Consider the location mixture generated by

$$X = Z + M$$

where $Z \sim N(0,1)$, and M has some distribution f_M with finite moments, with Z and M independent. This is equivalent to the location mixture

$$\begin{aligned} X|M = m &\sim \phi(x-m) \\ M &\sim f_M(m) \end{aligned}$$

where ϕ is the standard normal pdf. Then $E_{f_X}[X] = E_{f_M}[M] = \mu_M$, say, and

$$\begin{aligned} \varsigma &= E_{f_X}[(X-\mu)^3] = E_{f_X}[(X-\mu_M)^3] \\ &= E_{f_{X,M}}[(X-M+M-\mu_M)^3] \\ &= E_{f_{X,M}}[(X-M)^3 + 3(X-M)^2(M-\mu_M) + 3(X-M)(M-\mu_M)^2 + (M-\mu_M)^3] \\ &\equiv E_{f_Z}[Z^3] + 3E_{f_Z}[Z^2]E_{f_M}[(M-\mu_M)] + 3E_{f_Z}[Z]E_{f_M}[(M-\mu_M)^2] + E_{f_M}[(M-\mu_M)^3] \\ &= 0 + 3(1\times 0) + 3(0 \times Var_{f_M}[M]) + E_{f_M}[(M-\mu_M)^3] \\ &= E_{f_M}[(M-\mu_M)^3] \end{aligned}$$

which is non-zero if $E_{f_M}[(M-\mu_M)^3] \neq 0$. Thus skewness in the distribution of M induces skewness in the distribution of X.

5 Compute by direct calculation, or using mgfs.

(i) $X \sim Bernoulli(\theta)$: we have

$$E_{f_X}[X] = (0 \times (1-\theta)) + (1 \times \theta) = \theta$$

Let $Z = X - \theta$. Then $P[Z = 1-\theta] = \theta$, $P[Z = -\theta] = 1-\theta$, so

$$\begin{aligned}
r = 1 &\quad : \quad E_{f_Z}[Z] = ((-\theta) \times (1-\theta)) + ((1-\theta) \times \theta) = 0 \\
r = 2 &\quad : \quad E_{f_Z}[Z^2] = ((-\theta)^2 \times (1-\theta)) + ((1-\theta)^2 \times \theta) = \theta^2(1-\theta) + \theta(1-\theta)^2 \\
r = 3 &\quad : \quad E_{f_Z}[Z^3] = ((-\theta)^3 \times (1-\theta)) + ((1-\theta)^3 \times \theta) = -\theta^3(1-\theta) + \theta(1-\theta)^3 \\
r = 4 &\quad : \quad E_{f_Z}[Z^4] = ((-\theta)^4 \times (1-\theta)) + ((1-\theta)^4 \times \theta) = \theta^4(1-\theta) + \theta(1-\theta)^4.
\end{aligned}$$

Thus for X, the skewness is

$$\varsigma = \frac{E_{f_X}[(X-\mu)^3]}{\sigma^3} = \frac{E_{f_X}[Z^3]}{\{E_{f_X}[Z^2]\}^{3/2}} = \frac{-\theta^3(1-\theta) + \theta(1-\theta)^3}{\{\theta^2(1-\theta) + \theta(1-\theta)^2\}^{3/2}} = \frac{(1-2\theta)}{\{\theta(1-\theta)\}^{1/2}}$$

and the kurtosis is

$$\kappa = \frac{E_{f_X}[(X-\mu)^4]}{\sigma^2} = \frac{E_{f_X}[Z^4]}{\{E_{f_X}[Z^2]\}^2} = \frac{\theta^4(1-\theta) + \theta(1-\theta)^4}{\{\theta^2(1-\theta) + \theta(1-\theta)^2\}^4} = \frac{\theta^3 + (1-\theta)^3}{\{\theta(1-\theta)\}^3}$$

(ii) If $X \sim Poisson(\lambda)$, $E_{f_X}[X] = Var_{f_X}[X] = \lambda$, and $M_X(t) = \exp\{\lambda(e^t - 1)\}$. Let $Z = X - \lambda$, so that

$$M_Z(t) = e^{-\lambda t} M_X(t) = \exp\{\lambda(e^t - t - 1)\}$$

and

$$M_Z(t) = \sum_{r=0}^{\infty} \frac{\lambda^r}{r!} \left(\frac{t^2}{2} + \frac{t^3}{6} + \frac{t^4}{24} + \cdots \right)^r = \sum_{r=0}^{\infty} \frac{a_r}{r!} t^r.$$

We require the coefficients a_r for $r = 1, 2, 3$ and 4. Sufficient terms in the expansion can be obtained from the expansion of the first two terms

$$\lambda \left(\frac{t^2}{2} + \frac{t^3}{6} + \frac{t^4}{24} \right) + \frac{\lambda^2}{2} \left(\frac{t^2}{2} + \frac{t^3}{6} + \frac{t^4}{24} \right)^2 = \lambda \left(\frac{t^2}{2} + \frac{t^3}{6} + \frac{t^4}{24} \right) + \frac{\lambda^2}{2} \left(\frac{t^4}{4} + \frac{t^5}{6} + \frac{t^6}{36} + \cdots \right)$$

a_1 : No term in t in the expansion. Thus $E_{f_Z}[Z] = 0$.

a_2 : Term in t^2 in the expansion has coefficient $\lambda/2$. Thus $E_{f_Z}[Z^2] = \lambda$.

a_3 : Term in t^3 in the expansion has coefficient $\lambda/6$. Thus $E_{f_Z}[Z^3] = \lambda$.

a_4 : Term in t^4 in the expansion has coefficient $\lambda/24 + \lambda^2/8$. Thus $E_{f_Z}[Z^4] = \lambda + 3\lambda^2$.

Thus for X, the skewness is

$$\varsigma = \frac{E_{f_X}[(X-\mu)^3]}{\sigma^3} = \frac{E_{f_Z}[Z^3]}{\{E_{f_Z}[Z^2]\}^{3/2}} = \frac{\lambda}{\lambda^{3/2}} = \lambda^{-1/2}$$

and the kurtosis is

$$\kappa = \frac{E_{f_X}[(X-\mu)^4]}{\sigma^2} = \frac{E_{f_Z}[Z^4]}{\{E_{f_Z}[Z^2]\}^2} == \frac{\lambda + 3\lambda^2}{\lambda^2} = 3 + \frac{1}{\lambda}$$

(iii) If $X \sim Gamma(\alpha, \beta)$, we consider the standard case $Y \sim Gamma(\alpha, 1)$, and deduce the results from the relationship $X = Y/\beta$. Now,

$$M_Y(t) = \left(\frac{1}{1-t}\right)^{\alpha} = 1+\alpha t+\frac{\alpha(\alpha+1)}{2}t^2+\frac{\alpha(\alpha+1)(\alpha+2)}{6}t^3+\frac{\alpha(\alpha+1)(\alpha+2)(\alpha+3)}{24}t^4+\cdots$$

so if $Z = Y - E_{f_Y}[Y] = Y - \alpha$, then

$$M_Z(t) = e^{-\alpha t}\left[1+\alpha t+\frac{\alpha(\alpha+1)}{2}t^2+\frac{\alpha(\alpha+1)(\alpha+2)}{6}t^3+\frac{\alpha(\alpha+1)(\alpha+2)(\alpha+3)}{24}t^4+\cdots\right].$$

We have the results by expanding the exponential, multiplying and collecting terms. We need the first four terms, so it is sufficient to expand the exponential to

$$1-\alpha t+\frac{\alpha^2}{2}t^2-\frac{\alpha^3}{6}t^3+\frac{\alpha^4}{24}t^4$$

and multiply out. Thus

a_1 : No term in t in the expansion, as all cancel. Thus $E_{f_Z}[Z] = 0$.

a_2 : Term in t^2 in the expansion has coefficient

$$\frac{\alpha(\alpha+1)}{2}-\alpha^2+\frac{\alpha^2}{2}=\frac{\alpha}{2}.$$

Thus $E_{f_Z}[Z^2] = \alpha$.

a_3 : Term in t^3 in the expansion has coefficient

$$\frac{\alpha(\alpha+1)(\alpha+2)}{6}-\frac{\alpha^2(\alpha+1)}{2}+\frac{\alpha^3}{2}-\frac{\alpha^3}{6}=\frac{\alpha}{3}.$$

Thus $E_{f_Z}[Z^3] = 2\alpha$.

a_4 : Term in t^4 in the expansion has coefficient

$$\frac{\alpha(\alpha+1)(\alpha+2)(\alpha+3)}{24}-\frac{\alpha^2(\alpha+1)(\alpha+2)}{6}+\frac{\alpha^3(\alpha+1)}{4}-\frac{\alpha^4}{6}+\frac{\alpha^4}{24}=\frac{3\alpha(\alpha+2)}{24}$$

Thus $E_{f_Z}[Z^4] = 3\alpha(\alpha+2)$.

Hence

$$\begin{aligned} E_{f_Z}[Z^2] &= \alpha \\ E_{f_Z}[Z^3] &= 2\alpha \\ E_{f_Z}[Z^4] &= 3\alpha(\alpha+2) \end{aligned}$$

and thus for Y, the skewness is

$$\varsigma = \frac{E_{f_Y}[(Y-\alpha)^3]}{\sigma^3} = \frac{E_{f_Z}[Z^3]}{\{E_{f_Z}[Z^2]\}^{3/2}} = \frac{2\alpha}{\alpha^{3/2}} = \frac{2}{\alpha^{1/2}}$$

and the kurtosis is

$$\kappa = \frac{E_{f_Y}[(Y-\alpha)^4]}{\sigma^2} = \frac{E_{f_Z}[Z^4]}{\{E_{f_Z}[Z^2]\}^2} = \frac{3\alpha(\alpha+2)}{\alpha^2} = \frac{3(\alpha+2)}{\alpha} = 3+\frac{6}{\alpha}$$

Now

$$E_{f_Y}[(Y-\alpha)^r] = \beta^r E_{f_Y}[(X-\alpha/\beta)^r] = \beta^r E_{f_Y}[(X-\mu)^r]$$

where μ is the expectation of X. Thus the skewness and kurtosis of X are identical to the skewness and kurtosis of Y, as the factors involving β cancel.