ST202/ST206 – Autumn Term

Solutions to problem set 5

1. (a) $X \sim \text{Bernoulli}(p)$

$$\begin{aligned} M_X(t) &= E(e^{tX}) = \sum_x e^{tx} f_X(x) \\ &= e^{t0}P(X=0) + e^{t1}P(X=1) = (1-p) + pe^t \\ K_X(t) &= \log M_X(t) = \log[(1-p) + pe^t]. \end{aligned}$$

   (b) $Y \sim \text{Bin}(n, p)$

$$\begin{aligned} M_Y(t) &= E(e^{tY}) = \sum_y e^{ty} f_Y(y) \\ &= \sum_{y=0}^{n} e^{ty} \binom{n}{y} p^y (1-p)^{n-y} \\ &= \sum_{y=0}^{n} \binom{n}{y} (pe^t)^y (1-p)^{n-y} \\ &= [(1-p) + pe^t]^n \qquad \text{(binomial expansion)} \\ K_Y(t) &= \log M_Y(t) = n \log[(1-p) + pe^t]. \end{aligned}$$

   Note that if $X \sim \text{Bernoulli}(p)$ and $Y \sim \text{Bin}(n, p)$, then

$$\begin{aligned} M_Y(t) &= (M_X(t))^n, \\ K_Y(t) &= nK_X(t). \end{aligned}$$

   This is no coincidence; a $\text{Bin}(n, p)$ random variables can be viewed as the sum of n independent $\text{Bernoulli}(p)$ random variables.

   (c) $X \sim \text{Geometric}(p)$

$$\begin{aligned} M_X(t) &= E(e^{tX}) = \sum_x e^{tx} f_X(x) \\ &= \sum_{x=1}^{\infty} e^{tx} (1-p)^{x-1} p \\ &= pe^t \sum_{x=1}^{\infty} \left[(1-p)e^t\right]^{x-1} \qquad \text{(form a geometric series)} \\ &= \frac{pe^t}{1-(1-p)e^t} \qquad \text{(provided } |(1-p)e^t| < 1\text{),} \end{aligned}$$

which is well-defined for $|t| < -\log(1-p)$. The CGF is

$$K_X(t) = \log \frac{pe^t}{1-(1-p)e^t}.$$

(d) $Y \sim \text{NegBin}(r, p)$

$$\begin{aligned}
M_Y(t) = E(e^{tY}) &= \sum_y e^{ty} f_Y(y) \\
&= \sum_{x=r}^{\infty} e^{tx} \binom{x-1}{r-1} p^r (1-p)^{x-r} \\
&= (pe^t)^r \sum_{x=r}^{\infty} \frac{(x-1)!}{(r-1)!(x-r)!} \left[(1-p)e^t\right]^{x-r} \\
&= (pe^t)^r \sum_{y=0}^{\infty} \frac{(y+r-1)!}{(r-1)!y!} \left[(1-p)e^t\right]^{y} && \text{(let } y = x - r\text{)} \\
&= (pe^t)^r \left(1-(1-p)e^t\right)^{-r} && \text{(neg. bin. formula)} \\
&= \left(\frac{pe^t}{1-(1-p)e^t}\right)^r,
\end{aligned}$$

where, as before, we need $|(1-p)e^t| < 1$. The CGF is

$$K_Y(t) = r \log \frac{pe^t}{1-(1-p)e^t}.$$

Notice that, if $X \sim \text{Geometric}(p)$, we have

$$\begin{aligned}
M_Y(t) &= (M_X(t))^n, \\
K_Y(t) &= nK_X(t).
\end{aligned}$$

This relationship stems from the fact that the sum of r independent Geometric(p) random variables follows a NegBin(r, p) distribution.

2. In this question we work directly from the representation of the MGF as a polynomial,

$$M_X(t) = 1 + \mu_1' t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \dots,$$

and use the expansion

$$\log(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \dots$$

The cumulant generating function is then

$$\begin{aligned}K_X(t) = \log M_X(t) &= \log\left\{1 + \left(\mu_1' t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \dots\right)\right\} \\ &= \left(\mu_1' t + \mu_2' \frac{t^2}{2!} + \mu_3' \frac{t^3}{3!} + \dots\right) - \frac{1}{2}\left(\quad\right)^2 + \frac{1}{3}\left(\quad\right)^3 - \dots\end{aligned}$$

We need to find the coefficient of t^3 in the above equation. There are three terms in t^3, one from each of the first three sets of brackets in the expansion. Their sum is

$$\mu_3' \frac{t^3}{3!} + \frac{1}{2} 2\mu_1' t \mu_2' \frac{t^2}{2!} + \frac{1}{3}\left(\mu_1' t\right)^3 = \frac{t^3}{3!}(\mu_3' - 3\mu_1'\mu_2' + 2{\mu_1'}^3)\,,$$

where the coefficient of $t^3/3!$ in the final expression is equal to μ_3, the third central moment (standard result). As this coefficient in the expansion of the CGF is the third cumulant (by definition), we conclude that $\kappa_3 = \mu_3$.

3. The MGF for a standard normal random variable is

$$\begin{aligned}M_Z(t) &= \int_{-\infty}^{\infty} e^{zt} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(z^2 - 2tz)/2} dz \\ &= e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-(z^2 - 2tz + t^2)/2} dz = e^{t^2/2}\,, \text{ for } t \in \mathbb{R}.\end{aligned}$$

The last integral is equal to 1 because the integrand is a $\mathrm{N}(t, 1)$ density function. The MGF of X is then

$$\begin{aligned}M_X(t) &= \mathbb{E}\left[e^{tX}\right] = \mathbb{E}\left[e^{t(\mu+\sigma Z)}\right] \\ &= e^{\mu t}\,\mathbb{E}\left[e^{\sigma t Z}\right] = e^{\mu t} M_Z(\sigma t) = e^{\mu t + \sigma^2 t^2/2}\,, \text{ for } t \in \mathbb{R}.\end{aligned}$$

The mean of Y is, thus,

$$\mathbb{E}[Y] = \mathbb{E}[e^X] = M_X(1) = e^{\mu + \sigma^2/2}\,.$$

For the variance, we first work out

$$\mathbb{E}[Y^2] = M_X(2) = e^{2\mu + 4\sigma^2/2}\,,$$

so we have

$$\mathrm{Var}(Y) = e^{2\mu + 4\sigma^2/2} - (e^{\mu + \sigma^2/2})^2 = e^{2(\mu + \sigma^2/2)}(e^{\sigma^2} - 1).$$

4. We have

$$\begin{aligned}
M_X(t) = \mathbb{E}(e^{tX}) &= \int_{-\infty}^{\infty} e^{tx}\frac{1}{2}e^{-|x|}dx \qquad\qquad \text{(split into two)}\\
&= \frac{1}{2}\left(\int_{-\infty}^{0} e^{tx}e^{x}dx + \int_{0}^{\infty} e^{tx}e^{-x}dx\right)\\
&= \frac{1}{2}\left(\int_{-\infty}^{0} e^{(1+t)x}dx + \int_{0}^{\infty} e^{-(1-t)x}dx\right)\\
&= \frac{1}{2}\left(\left[\frac{1}{1+t}e^{(1+t)x}\right]_{-\infty}^{0} - \left[\frac{1}{1-t}e^{-(1-t)x}\right]_{0}^{\infty}\right)\\
&= \frac{1}{2}\left(\frac{1}{1+t} + \frac{1}{1-t}\right) = \frac{1}{1-t^2}\,,
\end{aligned}$$

assuming that $1+t > 0$ and $1-t > 0$, so that both integrals converge. This means that the MGF is defined for $|t| < 1$.

To work out the cumulants, we use the power series expansion of $\log(1-a)$,

$$\begin{aligned}
K_X(t) = -\log(1-t^2) &= t^2 + \frac{(t^2)^2}{2} + \frac{(t^2)^3}{3} + \dots\\
&= 0t + 2\,\frac{t^2}{2!} + 0\,\frac{t^3}{3!} + 12\,\frac{t^4}{4!} + \dots
\end{aligned}$$

We deduce that $\kappa_1 = 0$, $\kappa_2 = 2$, $\kappa_3 = 0$, and $\kappa_4 = 12$.

5. Let Z be a random variable with density

$$f_Z(z) = \frac{1}{k}, \quad \text{for} \quad -b < z < b.$$

   (a) The density must integrate to 1 over the support,

   $$1 = \int_{-\infty}^{\infty} f_Z(z)dz = \int_{-b}^{b} \frac{1}{k}dz = \frac{2b}{k}\,,$$

   so $k = 2b$.

   (b) We have

   $$\mathbb{E}(e^{tZ}) = \int_{-b}^{b} e^{tz}\frac{1}{2b}dz = \left[\frac{e^{tZ}}{2bt}\right]_{-b}^{b} = \frac{e^{tb} - e^{-tb}}{2bt}\,.$$

   (c) The MGF is not well-defined at $t = 0$, but we can use L'Hôpital's rule to find its limit,

   $$\lim_{t\to 0} \mathbb{E}(e^{tZ}) = \lim_{t\to 0} \frac{e^{tb} - e^{-tb}}{2bt} = \lim_{t\to 0} \frac{be^{tb} + be^{-tb}}{2b} = 1\,.$$

We can thus define

$$M_Z(t) = \begin{cases} \frac{e^{tb} - e^{-tb}}{2bt} & \text{if } t \neq 0, \\ 1 & \text{if } t = 0, \end{cases}$$

which is well-defined.