Solutions to Bonus Problems

Problem 1 (Moment generating functions). For the random variable X, the moments can be expressed as

$$E\left(X^n\right) = \frac{2^n}{n+1}, \; n = 1, 2, 3, \ldots$$

Find some (in fact, the unique) distribution of X having these moments.
Hint: Study the moment generating function of X and use the fact that

$$e^{tX} = \sum_{n=0}^{\infty} \frac{t^n X^n}{n!}$$

Solution. Following the hint, it follows that

$$M_X\left(t\right) = E\left(e^{tX}\right) = E\left(\sum_{n=0}^{\infty} \frac{t^n X^n}{n!}\right) = \sum_{n=0}^{\infty} \frac{t^n}{n!} E\left(X^n\right)$$

Next we use the fact that the moments of X are known.

$$M_X\left(t\right) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \cdot \frac{2^n}{n+1} = \sum_{n=0}^{\infty} \frac{\left(2t\right)^n}{\left(n+1\right)!}$$

We notice that the summand resembles the power series expansion of e^{2t}. We must, however, first make some adjustments. First we let $m = n+1$, i.e.

$$M_X\left(t\right) = \sum_{m=1}^{\infty} \frac{\left(2t\right)^{m-1}}{m!} = \frac{1}{2t} \sum_{m=1}^{\infty} \frac{\left(2t\right)^m}{m!}$$

Now the summand is correct, but because of the change of variable the summation starts at $m = 1$ and not at $m = 0$ as it should. We notice that the term for $m = 0$ is $\left(2t\right)^0 / 0! = 1$, and therefore

$$\sum_{m=1}^{\infty} \frac{\left(2t\right)^m}{m!} = \sum_{m=0}^{\infty} \frac{\left(2t\right)^m}{m!} - 1$$

And so, finally,

$$M_X\left(t\right) = \frac{1}{2t}\left(\sum_{m=0}^{\infty} \frac{\left(2t\right)^m}{m!} - 1\right) = \frac{e^{2t} - 1}{2t}$$

which we recognize as the mgf of $U\left(0, 2\right)$.

Problem 2 (Mgf for the $\chi^2(1)$-distribution). Let $Z \sim N(0,1)$ and consider $X = Z^2$. In the textbook, Section 2.3 Example C, the cdf-method is used to show that $X \sim \chi^2(1)$, i.e. the so called chi-square distribution with 1 degree of freedom. The general chi-square distribution with n degrees of freedom is actually a member of the gamma family where $\alpha = n/2$ and $\lambda = 1/2$, i.e. the $\chi^2(n)$-distribution is identical to the $Ga(n/2, 1/2)$-distribution. Even though it is not a linear transformation, the mgf method can be used to identify the distribution of $X = Z^2$. Determine

$$M_X(t) = E\left(e^{tX}\right) = \cdots$$

and compare this expression to the mgf of $\chi^2(1)$ given in the information about the common probability distributions.

Solution. By the definition of mgf,

$$M_X(t) = E\left(e^{tX}\right) = E\left(e^{tZ^2}\right) = \int_{-\infty}^{\infty} e^{tz^2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2(1-2t)} dz.$$

The integrand show similarities with (another) normal pdf, and to see which one it is we rewrite the expression as follows:

$$M_X(t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2(1-2t)} dz = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\frac{z^2}{\left(1/\sqrt{1-2t}\right)^2}} dz,$$

and it is clear that the integrand is similar to the pdf of $N(0, 1/(1-2t))$. It therefore follows that

$$M_X(t) = \frac{1}{\sqrt{1-2t}} \int_{-\infty}^{\infty} \frac{1}{\frac{1}{\sqrt{1-2t}} \cdot \sqrt{2\pi}} e^{-\frac{1}{2}\frac{z^2}{\left(1/\sqrt{1-2t}\right)^2}} dz = \frac{1}{\sqrt{1-2t}}, \quad t < \frac{1}{2},$$

which we recognize as the mgf of the $\chi^2(1)$-distribution.

Problem 3 (Covariance of linear combinations). We roll a fair die twice. Let X represent the sum of pips on the two rolls, and let Y represent the number of pips on the first roll minus the number of pips on the second roll. Compute $Cov\,(X, Y)$.

Hint. Let W_1 represent the number of pips on the first roll, and let W_2 represent the number of pips on the second roll. Now express X and Y in terms of W_1 and W_2, and then use formulas associated with the covariance of linear combinations.

Solution. We begin by making two important observations. (i) W_1 and W_2 are i.i.d random variables, and (ii) $X = W_1 + W_2$, $Y = W_1 - W_2$. The results concerning the covariance of linear combinations then gives us that

$$\begin{aligned} Cov\,(X, Y) &= Cov\,(W_1 + W_2, W_1 - W_2) = \\ &= Cov\,(W_1, W_2) + Cov\,(W_1, W_1) + Cov\,(W_2, W_1) - Cov\,(W_2, W_2) \end{aligned}$$

Since W_1 and W_2 are independent it follows that $Cov\,(W_1, W_2) = Cov\,(W_2, W_1) = 0$, and since they are identically distributed we have that $Cov\,(W_1, W_1) = Var\,(W_1) = Var\,(W_2) = Cov\,(W_2, W_2)$. We therefore conclude that

$$Cov\,(X, Y) = Var\,(W_1) - Var\,(W_2) = \mathbf{0}$$

As an alternative solution we can display the joint probability distribution of X and Y in a cross table as follows:

							x					
		2	3	4	5	6	7	8	9	10	11	12
	-5						1/6					
	-4					1/6		1/6				
	-3				1/6		1/6		1/6			
	-2			1/6		1/6		1/6		1/6		
	-1		1/6		1/6		1/6		1/6		1/6	
y	0	1/6		1/6		1/6		1/6		1/6		1/6
	1		1/6		1/6		1/6		1/6		1/6	
	2			1/6		1/6		1/6		1/6		
	3				1/6		1/6		1/6			
	4					1/6		1/6				
	5						1/6					

By symmetry, we must have that $E\,(Y) = 0$, and, also by symmetry, $E\,(XY) = 0$. This means that

$$Cov\,(X, Y) = E\,(XY) - E\,(X)\,E\,(Y) = \mathbf{0}$$

Problem 4 (Convergence in distribution). Let $X_p \sim NegBin\,(n, p)$. Show that pX_p converges in distribution as $p \to 0$ and determine the limiting distribution.

Hint. Determine the mgf of pX_p and then use the MacLaurin Series representation of the exponential function. Also, the limiting distribution is continuous.

Solution. Consider $\{Y_p = pX_p,\, 0 < p < 1\}$, i.e. a collection of random variables such that $X_p \sim NegBin(n, p)$. We have to find the mgf of Y_p and see what happens to this expression as $p \to 0$. First, the mgf of X_p is given by

$$M_{X_p}(t) = \left(\frac{pe^t}{1-(1-p)\,e^t}\right)^n, \quad t < -\ln(1-p)\,.$$

Since Y_p is a simple linear transformation of X_p it follows that

$$M_{Y_p}(t) = M_{X_p}(pt) = \left(\frac{pe^{pt}}{1-(1-p)\,e^{pt}}\right)^n = \left(\frac{p}{e^{-pt}-(1-p)}\right)^n, \quad t < -\frac{\ln(1-p)}{p},$$

where we in the last step divide both the numerator and the denominator by e^{pt} in order to simplify the expression. Both the numerator and the denominator will tend to zero as $p \to 0$ which means that no limit is obvious. In order to see what happens in the limit, a standard technique often used in probability theory is to replace the involved functions with their power series representations. In this case we need the power series representation of e^{-pt}, i.e.

$$e^{-pt} = \sum_{k=0}^{\infty} \frac{(-pt)^k}{k!} = 1 - pt + \frac{p^2t^2}{2!} - \frac{p^3t^3}{3!} + \cdots$$

Now

$$M_{Y_p}(t) = \left(\frac{p}{\left(1-pt+\frac{p^2t^2}{2!}-\frac{p^3t^3}{3!}+\cdots\right)-(1-p)}\right)^n = \left(\frac{p}{p-pt+\frac{p^2t^2}{2!}-\frac{p^3t^3}{3!}+\cdots}\right)^n.$$

Dividing both the numerator and the denominator by p results in the expression

$$M_{Y_p}(t) = \left(\frac{1}{1-t+\frac{pt^2}{2!}-\frac{p^2t^3}{3!}+\cdots}\right)^n \to \left(\frac{1}{1-t}\right)^n = (1-t)^{-n}, \quad \text{as} \quad p \to 0,$$

which we recognize as the mgf of $Ga(n, 1)$. It is therefore clear that $Y_p \xrightarrow{d} Ga(n, 1)$ as $p \to 0$.