Joint moment generating functions

Márton Balázs* and Bálint Tóth*

October 19, 2014

This little write-up is part of important foundations of probability that were left out of the unit Probability 1 due to lack of time and prerequisites. We have seen in Probability 1 how moment generating functions work for a single random variable, here we define it for joint distributions.

Definition 1 *Let $X_1,\, X_2,\, \dots,\, X_n$ be random variables. Their* joint moment generating function *is*

$$M(t_1,\, t_2,\, \dots,\, t_n) := \mathbf{E}\mathrm{e}^{t_1X_1+t_2X_2+\dots+t_nX_n} = \mathbf{E}\mathrm{e}^{\underline{t}^T\cdot\underline{X}}$$

(using a vector notation at the end).

The marginal moment generating functions are contained in a trivial manner:

$$M_{X_i}(t_i) = \mathbf{E}\mathrm{e}^{t_iX_i} = \mathbf{E}\mathrm{e}^{0X_1+0X_2+\dots+0X_{i-1}+t_iX_i+0X_{i+1}+\dots+0X_n} = M(0,\, 0,\, \dots,\, 0,\, t_i,\, 0,\, \dots,\, 0).$$

Uniqueness as seen for a single variable holds here too: knowing the joint moment generating function on an open neighbourhood of $\underline{0}$ uniquely determines the joint distribution. As a consequence, we have

Proposition 2 *The random variables $X_1,\, X_2,\, \dots,\, X_n$ are independent if and only if their joint moment generating function factorises around an open neighbourhood of zero:*

$$M(t_1,\, t_2,\, \dots,\, t_n) = M_{X_1}(t_1)\cdot M_{X_2}(t_2)\cdots M_{X_n}(t_n).$$

As an illustration we repeat an example from Probability 1 with moment generating functions.

Example 3 (marking of the Poisson distribution) *Suppose that a Poisson(λ) number of people enter the post office in the first hour in the morning, and each person is, independently of everything, female with probability p and male with probability $1-p$. We show that the number of females in the first hour is $X\sim$ Poi(λp), the number of males is $Y\sim$ Poi($\lambda(1-p)$) and these variables are independent.*

Based on the information given, we know that $X+Y\sim\mathrm{Poi}(\lambda)$; $(X\,|\,X+Y)\sim\mathrm{Binom}(X+Y,\, p)$. Therefore the joint moment generating function can be calculated via the tower rule:

$$M(t,\, s) = \mathbf{E}\big(\mathrm{e}^{tX+sY}\big) = \mathbf{E}\mathbf{E}\big(\mathrm{e}^{tX+sY}\,|\,X+Y\big) = \mathbf{E}\mathbf{E}\big(\mathrm{e}^{s(X+Y)}\mathrm{e}^{(t-s)X}\,|\,X+Y\big) = \mathbf{E}\big[\mathrm{e}^{s(X+Y)}\mathbf{E}\big(\mathrm{e}^{(t-s)X}\,|\,X+Y\big)\big],$$

since $X+Y$ is deterministic under conditioning on the value of $X+Y$. As the conditional distribution of X is Binomial, we substitute the Binomial moment generating function from Probability 1:

$$\mathbf{E}\big(\mathrm{e}^{(t-s)X}\,|\,X+Y\big) = M_{\mathrm{Binom}(X+Y,\,p)}(t-s) = \big(\mathrm{e}^{t-s}p+1-p\big)^{X+Y} = \mathrm{e}^{\ln(\mathrm{e}^{t-s}p+1-p)\cdot(X+Y)}.$$

With this we have

$$M(t,\, s) = \mathbf{E}\big[\mathrm{e}^{[s+\ln(\mathrm{e}^{t-s}p+1-p)]\cdot(X+Y)}\big].$$

Notice that this is just the moment generating function of $X+Y$ taken at $[s+\ln(\mathrm{e}^{t-s}p+1-p)]$. Now we use the fact that $X+Y\sim\mathrm{Poi}(\lambda)$ and the Poisson moment generating function from Probability 1:

$$M(t,\, s) = M_{\mathrm{Poi}(\lambda)}\big(s+\ln(\mathrm{e}^{t-s}p+1-p)\big) = \exp\big(\lambda(\mathrm{e}^{s+\ln(\mathrm{e}^{t-s}p+1-p)}-1)\big) =$$
$$= \mathrm{e}^{\lambda[\mathrm{e}^s(\mathrm{e}^{t-s}p+1-p)-1]} = \mathrm{e}^{\lambda p(\mathrm{e}^t-1)}\cdot\mathrm{e}^{\lambda(1-p)(\mathrm{e}^s-1)} = M_{\mathrm{Poi}(\lambda p)}(t)\cdot M_{\mathrm{Poi}(\lambda(1-p))}(s).$$

What we see here is the product of a Poi(λp) and a Poi($\lambda(1-p)$) moment generating function. By the uniqueness of the joint moment generating functions this implies that the joint distribution of X and Y is independent Poi(λp) and Poi($\lambda(1-p)$), respectively.

*University of Bristol / Budapest University of Technology and Economics