Probability Club 2!!!!

Peter Djemal

February 4, 2025

Moment Generating Functions

Given a random variable (RV) X with distribution f_X, we define the Moment Generating Function (MGF), presuming it converges, as:

$$M_X(t) = \mathbb{E}(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f_X \, \mathrm{d}x$$

It serves two important uses. The first is in calculating moments i.e. $\mathbb{E}(X)$, $\mathbb{E}(X^2)$ etc.

(a) Show that, unsurprisingly:

$$\mathbb{E}(X^n) = \left.\frac{\mathrm{d}^n M_X}{\mathrm{d}t^n}\right|_{t=0}$$

Find $M_X(t)$ for $X \sim \mathrm{Po}(\lambda)$ the Poisson distribution (the integral is replaced by a sum!). Use this to derive the first two moments ($n = 1, 2$) and the variance of X. Do the same for the standard normal $Y \sim N(0, 1)$.

$$\text{For reference } f_X(x) = \frac{\lambda^x}{x!} e^{-\lambda}, \ f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}$$

The second use is rather interesting. If X and Y are RVs with distributions f_X and f_Y, and if M_X and M_Y exist and are equal, then $f_X = f_Y$, i.e. each MGF represents a unique distribution. The proof of this is a nightmare so exercise caution in Googling it. It is important to note however that $\mathbb{E}(X^n) = \mathbb{E}(Y^n)\,\forall n$ would not be sufficient information to prove the distributions are identical, the MGFs must both exist (kind of, you can look up the details).

(b) Given the Chi-squared distribution of 1 degree of freedom, $Y \sim \chi_1^2$ with p.d.f.:

$$f_Y = \frac{e^{-y/2}}{\sqrt{2\pi y}} \ y > 0, \, 0 \text{ else}$$

derive:

$$M_Y(t) = (1 - 2t)^{-1/2}$$

Hint: the following will help and is well worth knowing in general:

$$\int_0^{\infty} x^{s-1} e^{-nx} \, \mathrm{d}x = \frac{\Gamma(s)}{n^s}, \ \Gamma(1/2) = \sqrt{\pi}$$

Given $X \sim N(0, 1)$, derive $M_{X^2}(t)$ and hence show that $X^2 \sim \chi_1^2$.

Proving the Central Limit Theorem using MGFs

Given a set of independent identically distributed (i.i.d.) X_i each with mean μ and **finite** variance σ^2, we are to show the Central Limit Theorem (CLT):

$$\lim_{n\to\infty}\left(\frac{1}{n}\sum_{k=1}^{n} X_k - \mu\right)\sqrt{n} \sim N(0,\sigma^2) \tag{*}$$

which implies that, for large n, the distribution of the sample mean $\bar{X}$ is roughly $N(\mu, \sigma^2/n)$.
We begin simplifying massively by defining the normalised $Y_i = (X_i - \mu)/\sigma$, then we consider the MGF of $\bar{Y}\sqrt{n}$ (the thing in the limit).
(c) Work this out, noting the i.i.d. conditions, you will get:

$$M_{\bar{Y}\sqrt{n}}(t) = \left[M_{Y_i}\left(\frac{t}{\sqrt{n}}\right)\right]^n$$

Taylor expand the M_{Y_i} and take the limit to derive the final MGF. Use this and results from previous questions to prove the CLT.

Bonus HARD and DIRTY problem - The Cauchy distribution

As an investigation of the ”finite variance” part of the CLT conditions, consider the Cauchy distribution $X \sim \text{Cauchy}(0,1)$ with p.d.f.:

$$f_X = \frac{1}{\pi}\frac{1}{1+x^2}$$

This effectively has $\sigma^2 = \infty$. We cannot use MGFs here since the integral wouldn't converge in $\mathbb{R}$, but it's about time we took a break from them anyway. We can use the fact that $X_1 + X_2$ has a distribution given by convolution. Hence for two Cauchy's:

$$f_{X_1+X_2}(t) = \frac{1}{\pi^2}\int_{-\infty}^{\infty}\frac{1}{1+x^2}\frac{1}{1+(t-x)^2}\,\mathrm{d}x$$

(d) Use any method e.g. partial fractions, complex analysis, to evaluate this integral. By repeating this pattern inductively to get $f_{X_1+\cdots+X_{2N}}(t)$, what is the p.d.f. $f_{\bar{X}_n}(x)$? What does the limit in the LHS of (*) evaluate to?

Hints and Solutions - ONLY WHEN NECESSARY

(a) For the Poisson distribution, you should get:

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

which will give

$$\mathbb{E}(X) = \lambda,\ \mathbb{E}(X^2) = \lambda(\lambda+1),\ \mathrm{Var}(X) = \lambda$$

For $N(0,1)$ similarly we have:

$$M_X(t) = e^{t^2/2},\ \mathbb{E}(X) = 0,\ \mathbb{E}(X^2) = 1,\ \mathrm{Var}(X) = 1$$

(b) Note that:

$$\mathbb{E}(e^{tX^2}) = \int_{-\infty}^{\infty} e^{tx^2} f_X \,\mathrm{d}x$$

Obviously you will get $M_{X^2} = M_Y$ which allows use of the stated theorem.

(c) Due to the i.i.d. conditions, $\mathbb{E}(e^{(t/\sqrt{n})\sum_i Y_i}) = \prod_i \mathbb{E}(e^{(t/\sqrt{n})Y_i})$. Given that each Y_i is normalised, $M_{Y_i}(t/\sqrt{n})^n = (1 + \frac{1}{2n}t^2 + \cdots)^n \to e^{t^2/2}$, which matches the MGF for $N(0,1)$ found in (a).

(d) You will get:

$$f_{X_1+X_2}(t) = \frac{1}{\pi}\frac{2}{x^2+4}$$

so a flattened out Cauchy, repeating this argument for $n = 2^N$:

$$f_{X_1+\cdots+X_{2^N}}(t) = \frac{1}{\pi}\frac{2^N}{t^2+4^N} \implies f_{\bar{X}_n}(x) = \frac{1}{\pi}\frac{1}{1+x^2}$$

the same as before! Note the "relative deviation" of $\bar{X}_n$ is constant, instead of decreasing by $1/\sqrt{n}$ in the case of a Gaussian. For the limit:

$$= \frac{1}{\pi}\frac{\sqrt{n}}{n+x^2} \to 0$$

hence the finite variance assumption is needed! (okay fine you can argue that the Gaussian of "infinite" variance would also be a zero function, so the limits indeed match. However! This shows that you at no point can *approximate* the distribution as normal)