Lecture Notes 8: Probability and Statistics

In the last lecture notes we have explored the MGF and discussed different properties of a MGF. One strong property of a MGF is that if it exists in a neighborhood of zero then it uniquely characterizes the distribution function. For example let us look at the MGF of a gamma distribution with the shape parameter $\alpha > 0$ and $\lambda > 0$. We will denote this as Gamma(α, λ). If the random variable X has Gamma(α, λ) distribution, then the PDF of X is

$$f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}; \quad x > 0, \tag{1}$$

and zero, otherwise. Therefore the moment generating function of X becomes

$$\begin{aligned} M_X(t) &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{tx} x^{\alpha-1} e^{-\lambda x} dx \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{tx} x^{\alpha-1} e^{-x(\lambda - t)} dx \\ &= \frac{\lambda^\alpha}{(\lambda - t)^\alpha} \quad \text{for} \quad |t| < \lambda. \end{aligned}$$

Hence, we can say that if a random variable Y has the MGF

$$M_Y(t) = \frac{\lambda^\alpha}{(\lambda - t)^\alpha}; \quad \text{for} \quad |t| < \lambda,$$

then Y has a gamma distribution with the shape parameter α and scale parameter λ. Now let us look at a simple application of this fact. Suppose X has a gamma distribution with PDF (1). Let us look at a new random variable $Y = 10X$, what will be the distribution function of Y. Let us look at the MGF of Y.

$$\begin{aligned} M_Y(t) &= E(e^{tY}) = E(e^{10tX}) = M_X(10t) = \frac{\lambda^\alpha}{(\lambda - 10t)^\alpha} \quad \text{for} \quad |10t| < \lambda \\ &= \frac{\lambda_1^\alpha}{(\lambda_1 - t)^\alpha} \quad \text{for} \quad |t| < \lambda_1 = \frac{\lambda}{10}. \end{aligned}$$

Therefore, we can say that Y has a gamma distribution with the shape parameter α and scale parameter $\lambda_1 = \dfrac{\lambda}{10}$. We will see several such applications of the MGF.

Problem: Suppose X is a N(0,1) random variable, i.e. a normal distribution with mean zero and variance 1. Let $Y = X^2$. Find the moment generating function of Y. Can you identify its distribution?

Further another important feature of the MGF is that if the MGF of a random variable X exists in a neighborhood of zero, then all the moments of X exists, and they can be obtained as follows. Let us look at

$$\begin{aligned} M_X'(t) &= \frac{d}{dt}M_X(t) = \frac{d}{dt}\int_{-\infty}^{\infty} e^{xt} f_X(x)dx = \int_{-\infty}^{\infty} \left\{\frac{d}{dt}e^{xt}\right\} f_X(x)dx \\ &= \int_{-\infty}^{\infty} xe^{xt} f_X(x)dx. \end{aligned}$$

Hence, $M_X'(0) = E(X)$. Similarly,

$$\begin{aligned} M_X''(t) &= \frac{d^2}{dt^2}M_X(t) = \frac{d^2}{dt^2}\int_{-\infty}^{\infty} e^{xt} f_X(x)dx = \int_{-\infty}^{\infty} \left\{\frac{d^2}{dt^2}e^{xt}\right\} f_X(x)dx \\ &= \int_{-\infty}^{\infty} x^2 e^{xt} f_X(x)dx. \end{aligned}$$

and $M_X''(0) = E(X^2)$. We can easily obtain $M_X^{(k)}(0) = E(X^k)$, where $M_X^{(k)}(t) = \dfrac{d^k}{dt^k}M_X(t)$, for $k = 1, 2, \ldots$.

Another related quantity is known as cumulant generating function (CGF), and it is defined as

$$K_X(t) = \ln M_X(t).$$

It is obvious that if $K_X(t)$ exists in a neighborhood of zero iff $M_X(t)$ exists in a neighborhood of zero. Now observe that

$$K_X'(t) = \frac{d}{dt}K_X(t) = \frac{M_X'(t)}{M_X(t)}.$$

Hence, $K_X'(0) = \dfrac{M_X'(0)}{M_X(0)} = E(X)$.

$$K_X''(t) = \frac{d^2}{dt^2}K_X(t) = \frac{M_X''(t)M_X(t) - (M_X'(t))^2}{(M_X(t))^2}.$$

Hence

$$K''_X(0) = E(X^2) - (E(X))^2 = E(X - E(X))^2 = V(X).$$

Now we will discuss several inequalities in probability which turns out to be quite useful in different applications. We will prove all the results for continuous case but they are true even for the discrete case also, unless it is stated explicitly otherwise. We have the following results.

Results: If X is a random variable with m-th order moment finite for $m > 0$, i.e.

$$\int_{-\infty}^{\infty} |x|^m f_X(x) dx < \infty,$$

then for any $0 < k < \infty$, k-th order moment is also finite. The proof is very simple as follows

$$\begin{aligned}
\int_{-\infty}^{\infty} |x|^k f_X(x) dx &= \int_{|x| \le 1} |x|^k f_X(x) dx + \int_{|x|>1} |x|^k f_X(x) dx \\
&\le \int_{|x| \le 1} f_X(x) dx + \int_{|x|>1} |x|^m f_X(x) dx < \infty.
\end{aligned}$$

Problem: If X is a bounded random variable i.e. there exists a $0 < K < \infty$, such that $|X| \le K$ with probability one. Then all the moments of X exist.

Problem: What do you think about the converse, i.e. if all the moments of a random variable exist is it always a bounded random variable?

Another important inequality is knows as the Chebyshev's inequality, and it is as follows.

Chebyshev's Inequality: If $u(x)$ is a non-negative function and for the random variable X, $E(u(X))$ exists. Then for any $c > 0$

$$P(u(X) \ge c) \le \frac{E(u(X))}{c}.$$

This turns out to be a very useful inequality. For example let us look at the following example. Suppose X is a gamma random variable with the following PDF

$$f_X(x) = \frac{2^{10}}{9!} x^9 e^{-2x}; \qquad x > 0,$$

and zero, otherwise. Here the shape parameter $\alpha = 10$ and $\lambda = 2$. Suppose we want to compute the following probability: $P(X > 2)$. It is

$$P(X > 2) = \int_2^\infty \frac{2^{10}}{9!} x^9 e^{-2x} dx.$$

Clearly it cannot be obtained in explicit forms. It has to be computed using numerical integration. On the other hand using the Chebyshev's inequality, we can get an upper bound of this probability, as follows. Let us take $u(x) = x$. Therefore

$$P(u(X) > 2) \le \frac{E(u(X)}{2} = \frac{1}{2}, \qquad \text{as} \qquad E(u(X)) = 1.$$

Let us look at another example. Suppose X is a $\mathrm{N}(0, 1/2)$, i.e. normal random variable with mean zero and variance 1/2. Therefore the PDF of X becomes

$$f_X(x) = \frac{\sqrt{2}}{\sqrt{2\pi}} e^{-x^2}; \qquad -\infty < x < \infty.$$

Suppose we want to compute the probability $P(|X| \le 3)$, and it will be

$$P(|X| \le 3) = \int_{-3}^{3} \frac{\sqrt{2}}{\sqrt{2\pi}} e^{-x^2} dx.$$

In this case also this probability cannot be obtained in explicit forms. But we will be able to get a lower bound based on Chebyshev's inequality. Let us take $u(x) = x^2$, therefore

$$P(u(X) \le c) \ge 1 - \frac{E(X^2)}{c}.$$

Hence, in this case it becomes

$$P(|X| \le 3) = P(X^2 \le 9) \ge 1 - \frac{E(X^2)}{9} = 1 - \frac{V(X)}{9} = 1 - \frac{1}{18} = \frac{17}{18}.$$

Now we are going to state another important inequality and it is known as Jensen's inequality. We need the following definition for that purpose. A real valued function $f(x)$ defined on an open interval I, is said to be a convex function if for any $x, y \in I$, and for any $0 < \alpha < 1$,

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y).$$

Further, the function is called strictly convex, if

$$f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha)f(y).$$

Some of the simple examples of the convex function are; $f(x) = x^2$, $f(x) = x^4$, for $x > 0$, $f(x) = -\ln(x)$ etc. One simple characteristic of the convex function is the following. We do not give the proof, it is available in any standard real analysis text book. If $f(x)$ is a real valued function whose second derivative exists in I, then $f(x)$ is convex if and only if $f''(x) \geq 0$, and it is strictly convex, if $f''(x) > 0$, for all $x \in I$. Now we state the Jensen's inequality. Suppose X is a random variable defined in an open set I, and $\phi(x)$ is convex function whose second derivative exists in I, then

$$\phi(E(X)) \leq E(\phi(X)), \tag{2}$$

provided all the expectation exists. Now to prove (2) let us denote $\mu = E(X)$. Now to expand $\phi(x)$ around μ, we obtain

$$\phi(x) = \phi(\mu) + (x-\mu)\phi'(\mu) + \frac{(x-\mu)^2}{2}\phi''(\tilde{x}),$$

here $\tilde{x}$ is a point on the line joining x and μ. Since $\phi''(\tilde{x}) \geq 0$,

$$\phi(x) \geq \phi(\mu) + (x-\mu)\phi'(\mu).$$

Therefore,

$$E(\phi(X)) \geq \phi(\mu) + E(X-\mu)\phi'(\mu) \quad \Rightarrow \quad E(\phi(X)) \geq \phi(E(X)).$$

Now if we take $\phi(x) = x^2$, it is a strict convex function. Therefore, if X is a random variable with finite second moment, then $(E(X))^2 < E(X^2)$.