

Lecture Notes: The Moment Generating Function

Probability for Statisticians - Lecture 2 Supplement

April 3, 2025

1 Random Variables: Unveiling the Moment Generating Function

Within the fascinating world of probability, we often seek tools to concisely describe the behavior of random variables. One such powerful instrument is the Moment Generating Function (MGF). As its name intriguingly suggests, it provides a systematic way to 'generate' the moments (like the mean, variance, etc.) of a random variable.

1.1 Definition: Capturing the Essence

Let's formally define this function.

Definition 1.1 (Moment Generating Function (MGF)). *Let X be a random variable with probability density function (PDF) $f(x)$ or probability mass function (PMF) $P(X = x_k)$. Suppose there exists some positive value $\delta > 0$ such that the expected value $\mathbb{E}[e^{\delta|X|}]$ is finite ($\mathbb{E}[e^{\delta|X|}] < \infty$). Then, the **Moment Generating Function (MGF)** of X , denoted by $M_X(t)$, is defined as:*

$$M_X(t) = \mathbb{E}[e^{tX}]$$

for all real numbers t such that $|t| < \delta$. Explicitly:

- Continuous case: $M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$
- Discrete case: $M_X(t) = \sum_k e^{tx_k} P(X = x_k)$

Remark 1.2 (The Existence Condition). *The condition $\mathbb{E}[e^{\delta|X|}] < \infty$ for some $\delta > 0$ is crucial. It ensures that the expectation defining the MGF, $\mathbb{E}[e^{tX}]$, converges (i.e., the integral or sum yields a finite value) within a certain open interval $(-\delta, \delta)$ around $t = 0$. This guarantees that $M_X(t)$ is a well-defined, finite function in a neighborhood of the origin, which is essential for the properties we'll explore next. If X is non-negative, the condition simplifies to $\mathbb{E}[e^{\delta X}] < \infty$.*

1.2 Lemma 1.5: Generating Moments Through Differentiation

Now, let's uncover why the MGF is aptly named. The following lemma reveals the connection between the MGF and the moments of the random variable X .

Lemma 1.3 (Moments from the MGF). *Suppose the Moment Generating Function $M_X(t)$ of a random variable X exists (i.e., is finite) in an open interval containing $t = 0$, say $(-\delta, \delta)$ for some $\delta > 0$. Then, all moments of X , denoted by $\mathbb{E}[X^p]$ for $p = 1, 2, 3, \dots$, exist and are finite. Furthermore, these moments can be obtained by differentiating the MGF $M_X(t)$ with respect to t and evaluating the result at $t = 0$:*

$$\mathbb{E}[X^p] = \left. \frac{d^p}{dt^p} M_X(t) \right|_{t=0}$$

for all positive integers $p \in \mathbb{N}$.

Proof of Lemma 1.3. We are given that $M_X(t) = \mathbb{E}[e^{tX}]$ is defined and finite on an open interval $(-\delta, \delta)$ for some $\delta > 0$. Let p be any positive integer.

Step 1: Show all moments $\mathbb{E}[|X|^p]$ are finite.

(This part remains the same as before) ... Consider the Taylor series expansion of the exponential function $e^u = \sum_{j=0}^{\infty} \frac{u^j}{j!}$. For any t_0 such that $0 < t_0 < \delta$, we can write:

$$e^{t_0|X|} = \sum_{j=0}^{\infty} \frac{(t_0|X|)^j}{j!} = 1 + t_0|X| + \frac{(t_0|X|)^2}{2!} + \dots + \frac{(t_0|X|)^p}{p!} + \dots$$

Since all terms in the series are non-negative, the sum is greater than or equal to any single term. Let's pick the term where $j = p$:

$$e^{t_0|X|} \geq \frac{(t_0|X|)^p}{p!} = \frac{t_0^p |X|^p}{p!}$$

Rearranging this inequality gives us:

$$|X|^p \leq \frac{p!}{t_0^p} e^{t_0|X|}$$

Now, let's take the expectation of both sides. Since $p!$ and t_0^p are constants for a fixed t_0 and p , we can pull them out of the expectation:

$$\mathbb{E}[|X|^p] \leq \mathbb{E}\left[\frac{p!}{t_0^p} e^{t_0|X|}\right] = \frac{p!}{t_0^p} \mathbb{E}[e^{t_0|X|}]$$

The initial assumption is that $\mathbb{E}[e^{\delta|X|}] < \infty$. Since $0 < t_0 < \delta$, it follows that $\mathbb{E}[e^{t_0|X|}]$ is also finite. (A brief justification: $e^{t_0|X|} \leq 1 + e^{\delta|X|}$ for all X , ensuring $\mathbb{E}[e^{t_0|X|}]$ is finite if $\mathbb{E}[e^{\delta|X|}]$ is).

Therefore, for our chosen t_0 (where $0 < t_0 < \delta$), the term $\mathbb{E}[e^{t_0|X|}]$ on the right-hand side is finite. Since $p!/t_0^p$ is also a finite constant, the entire right-hand side is finite. This forces the left-hand side, $\mathbb{E}[|X|^p]$, to be finite as well. Since this holds for any positive integer p , we conclude that all absolute moments, and hence all moments $\mathbb{E}[X^p]$, exist and are finite.

Remark 1.4 (Justification for Interchanging Differentiation and Expectation). *A crucial step in Step 2 below involves calculating $\frac{d}{dt} M_X(t)$ by bringing the derivative inside the expectation:*

$$\frac{d}{dt} \mathbb{E}[e^{tX}] \stackrel{?}{=} \mathbb{E}\left[\frac{d}{dt} e^{tX}\right] = \mathbb{E}[Xe^{tX}]$$

This interchange of differentiation and expectation (which is an integral or a sum)

requires careful justification. Let's delve into the discrete case first, as the argument often feels more accessible.

Discrete Case: Differentiating the Series Term-by-Term

Here, $M_X(t) = \sum_k e^{tx_k} P(X = x_k)$. We want to show that we can differentiate this series term by term:

$$\frac{d}{dt} \left(\sum_k e^{tx_k} P(X = x_k) \right) \stackrel{?}{=} \sum_k \frac{d}{dt} (e^{tx_k} P(X = x_k)) = \sum_k x_k e^{tx_k} P(X = x_k)$$

Calculus tells us this is permissible if the series of derivatives converges uniformly on the interval of interest. Let's consider any closed interval $[-\delta', \delta']$ inside our original interval $(-\delta, \delta)$, where $0 < \delta' < \delta$.

We can use the **Weierstrass M-Test** to establish uniform convergence. This test requires us to find a sequence of non-negative numbers M_k such that:

1. The absolute value of each term in the derivative series is bounded by M_k :

$$\left| \frac{d}{dt} (e^{tx_k} P(X = x_k)) \right| = |x_k e^{tx_k} P(X = x_k)| \leq M_k$$

for all $t \in [-\delta', \delta']$.

2. The series $\sum_k M_k$ converges.

Let's find a suitable M_k . For $t \in [-\delta', \delta']$, we have $|t| \leq \delta'$. Thus:

$$|x_k e^{tx_k} P(X = x_k)| = |x_k| e^{tx_k} P(X = x_k) \leq |x_k| e^{|\delta'| |x_k|} P(X = x_k) \leq |x_k| e^{\delta' |x_k|} P(X = x_k)$$

So, we can choose $M_k = |x_k| e^{\delta' |x_k|} P(X = x_k)$. Now, does $\sum_k M_k$ converge?

$$\sum_k M_k = \sum_k |x_k| e^{\delta' |x_k|} P(X = x_k) = \mathbb{E}[|X| e^{\delta' |X|}]$$

Is this expectation finite? Yes! This is the payoff from our initial assumption that $\mathbb{E}[e^{\delta |X|}]$ is finite for some $\delta > 0$. As we reasoned in Step 1, this finiteness implies $\mathbb{E}[|X|^p] < \infty$ for all p . It also guarantees that $\mathbb{E}[|X|^k e^{\delta' |X|}] < \infty$ for any integer $k \geq 0$ and any $0 < \delta' < \delta$. (The argument involves showing that $e^{\delta |X|}$ grows faster than any polynomial $|X|^k$, ensuring the combined expectation remains finite within $(-\delta, \delta)$). For $k = 1$, we have $\mathbb{E}[|X| e^{\delta' |X|}] < \infty$.

Since we found a convergent series $\sum_k M_k$ that bounds the terms of the derivative series for all $t \in [-\delta', \delta']$, the Weierstrass M-Test confirms that the series of derivatives $\sum_k x_k e^{tx_k} P(X = x_k)$ converges uniformly on $[-\delta', \delta']$. This justifies differentiating the original MGF series term-by-term within $(-\delta, \delta)$.

Continuous Case: Leibniz Integral Rule

An analogous argument holds for the continuous case, $M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$. Here, the justification relies on the **Leibniz Integral Rule**. The key step is again showing that the derivative term inside the integral, $x e^{tx} f(x)$, is dominated in absolute value by an integrable function for $t \in [-\delta', \delta']$. The dominating function is $H(x) = |x| e^{\delta' |x|} f(x)$, and its integrability, $\int H(x) dx = \mathbb{E}[|X| e^{\delta' |X|}] < \infty$, is guaranteed by the same fundamental condition $\mathbb{E}[e^{\delta |X|}] < \infty$.

Higher Derivatives

The same logic extends to higher derivatives. For the p -th derivative, we need to justify swapping $\frac{d^p}{dt^p}$ with the expectation. This requires showing that $\mathbb{E}[|X|^p e^{\delta'|X|}] < \infty$, which, as mentioned, also follows from the initial MGF existence condition. Therefore, the interchange is valid for all orders p .

Step 2: Relate derivatives of $M_X(t)$ to moments.

Now, let's differentiate $M_X(t) = \mathbb{E}[e^{tX}]$ with respect to t . Based on the rigorous justification provided in Remark 1.4 (whether viewing expectation as a sum or integral), we can swap the order of differentiation and expectation for $t \in (-\delta, \delta)$:

$$\frac{d}{dt}M_X(t) = \frac{d}{dt}\mathbb{E}[e^{tX}] = \mathbb{E}\left[\frac{d}{dt}e^{tX}\right] = \mathbb{E}[Xe^{tX}]$$

Differentiating again, the same justification applies:

$$\frac{d^2}{dt^2}M_X(t) = \frac{d}{dt}\mathbb{E}[Xe^{tX}] = \mathbb{E}\left[\frac{d}{dt}(Xe^{tX})\right] = \mathbb{E}[X^2e^{tX}]$$

Continuing this process p times, we arrive at:

$$\frac{d^p}{dt^p}M_X(t) = \mathbb{E}[X^p e^{tX}]$$

This holds for all $t \in (-\delta, \delta)$. Finally, we evaluate this p -th derivative at $t = 0$. Since $t = 0$ is within the interval $(-\delta, \delta)$, the expression is well-defined:

$$\left.\frac{d^p}{dt^p}M_X(t)\right|_{t=0} = \mathbb{E}[X^p e^{tX}]\Big|_{t=0} = \mathbb{E}[X^p e^{0 \cdot X}] = \mathbb{E}[X^p \cdot 1] = \mathbb{E}[X^p]$$

This completes the proof. □

Remark 1.5 (Why is this useful?). Lemma 1.5 provides a powerful analytical technique. If we can find the MGF of a random variable (perhaps by evaluating the expectation $\mathbb{E}[e^{tX}]$ directly from its probability density/mass function), we can then find *any* moment $\mathbb{E}[X^p]$ simply by repeated differentiation and evaluation at $t = 0$. This is often much easier than calculating $\mathbb{E}[X^p] = \int_{-\infty}^{\infty} x^p f(x) dx$ or $\mathbb{E}[X^p] = \sum_x x^p P(X = x)$ directly, especially for higher moments (large p).