> **Remark** (Why MGFs?)**.** Moment generating functions are the single most powerful tool in this course: with them we can *prove* that the sum of independent Poissons is Poisson, that the sum of independent normals is normal, and – later – the Central Limit Theorem itself.

Moments

Definition. The nth moment of a random variable X is $\mathbb{E}[X^n]$.

The first moment is the mean. The second moment gives the variance via $\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$. Higher moments measure finer features of the shape of a distribution: the third is related to *skewness* (asymmetry), the fourth to *kurtosis* (heaviness of the tails). The name comes from mechanics: $\mathbb{E}[X]$ is the centre of mass of the distribution, and the second moment about the mean is its moment of inertia.

It would be convenient to have a single object that stores *all* the moments at once. There is one.

The Moment Generating Function

Definition. The moment generating function (MGF) of a random variable X is the function

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

so for discrete and continuous variables respectively,

$$M_X(t) = \sum_x e^{tx}\,\mathbb{P}(X = x), \qquad M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\,\mathrm{d}x$$

> **Remark** (Convergence)**.** The sum or integral does not always converge: we require the MGF to exist (be finite) for all t in some interval around 0. For most of our distributions this is fine, but some heavy-tailed distributions (e.g. the Cauchy distribution) have no MGF at all. We will quietly assume everything converges where we need it to.

Why does this "generate moments"?

Theorem

$$M_X(t) = 1 + \mathbb{E}[X]t + \frac{\mathbb{E}[X^2]}{2!}t^2 + \frac{\mathbb{E}[X^3]}{3!}t^3 + \cdots$$

so the nth moment is $n!$ times the coefficient of t^n. Equivalently, differentiating n times and setting $t = 0$,

$$M_X^{(n)}(0) = \mathbb{E}[X^n]$$

The proof is simply to expand the exponential as a series.

In particular:

$$\mathbb{E}[X] = M_X'(0), \qquad \text{Var}[X] = M_X''(0) - (M_X'(0))^2$$

Example

A random variable X has MGF $M_X(t) = (1-2t)^{-3}$ for $t < \frac{1}{2}$. Find $\mathbb{E}[X]$ and $\text{Var}[X]$.

Properties of MGFs

Theorem (Scaling and shifting)
For constants a, b:

$$M_{aX}(t) = M_X(at), \qquad M_{X+b}(t) = e^{bt} M_X(t), \qquad M_{aX+b}(t) = e^{bt} M_X(at)$$

Theorem (Independence)
If X and Y are independent, then

$$M_{X+Y}(t) = M_X(t) M_Y(t)$$

This last property is the reason MGFs are so useful: *adding independent random variables multiplies their MGFs* – and multiplying functions is easy.

The final ingredient is that the MGF pins down the distribution:

Fact (Uniqueness) — If two random variables have the same MGF (finite on an interval around 0), then they have the same distribution. So if we compute $M_{X+Y}(t)$ and recognise it as the MGF of a known distribution, then $X + Y$ *has* that distribution. (The proof is well beyond this course.)

MGFs of the Standard Distributions

Bernoulli and binomial

Example

Find the MGF of $X \sim B(1, p)$ (Bernoulli), and hence of $Y \sim B(n, p)$.

Poisson

Example

Show that if $X \sim \text{Po}(\lambda)$ then $M_X(t) = e^{\lambda(e^t - 1)}$.

Example (Class practice)

Use this MGF to verify that $\mathbb{E}[X] = \lambda$ and $\text{Var}[X] = \lambda$.

Continuous uniform

Example

Show that if $X \sim U[a,b]$ then $M_X(t) = \dfrac{e^{bt} - e^{at}}{t(b-a)}$ for $t \neq 0$.

Exponential

Example

Show that if $X \sim \text{Exp}(\lambda)$ then $M_X(t) = \dfrac{\lambda}{\lambda - t}$ for $t < \lambda$, and use it to find $\mathbb{E}[X]$ and $\text{Var}[X]$.

Example (OCR S4, June 2013)

The continuous random variable X has probability density function given by

$$f(x) = \begin{cases} \frac{1}{4}xe^{-\frac{1}{2}x} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

(i) Show that the moment generating function of X is $(1-2t)^{-2}$ for $t < \frac{1}{2}$, and state why the condition $t < \frac{1}{2}$ is necessary.

(ii) Use the moment generating function to find $\text{Var}[X]$.

Normal

Example

Show that the standard normal $Z \sim N(0,1)$ has MGF $M_Z(t) = e^{t^2/2}$.

Example

Deduce the MGF of $X \sim N(\mu, \sigma^2)$.

Example (Class practice)

Find the MGF of the geometric distribution $X \sim \text{Geo}(p)$, where $\mathbb{P}(X = k) = q^{k-1}p$ for $k = 1, 2, 3, \ldots$ and $q = 1 - p$. For which t does it exist?

Example (OCR S4, June 2017)

The discrete random variable X is such that $\mathbb{P}(X = x) = \frac{3}{4}\left(\frac{1}{4}\right)^x$ for $x = 0, 1, 2, \ldots$

(i) Show that the moment generating function of X can be written as $M_X(t) = \dfrac{3}{4 - e^t}$.

(ii) Find the range of values of t for which the formula for $M_X(t)$ in part (i) is valid.

(iii) Use $M_X(t)$ to find $\mathbb{E}[X]$ and $\mathrm{Var}[X]$.

Fact (Summary of standard MGFs) —

Bernoulli $B(1,p)$	$q + pe^t$	
Binomial $B(n,p)$	$(q + pe^t)^n$	
Poisson $\mathrm{Po}(\lambda)$	$e^{\lambda(e^t - 1)}$	
Uniform $U[a,b]$	$\dfrac{e^{bt} - e^{at}}{t(b-a)}$	$t \neq 0$
Exponential $\mathrm{Exp}(\lambda)$	$\dfrac{\lambda}{\lambda - t}$	$t < \lambda$
Normal $N(\mu, \sigma^2)$	$e^{\mu t + \frac{1}{2}\sigma^2 t^2}$	

Applications: Sums of Independent Variables

Theorem (Sum of independent Poissons)

If $X \sim \text{Po}(\lambda)$ and $Y \sim \text{Po}(\mu)$ are independent, then $X + Y \sim \text{Po}(\lambda + \mu)$.

Example

Prove this using MGFs. (Compare the direct proof via convolution and the binomial theorem from the Poisson chapter – this one is three lines.)

Example (OCR S4, June 2011)

The discrete random variable X has moment generating function $\left(\frac{1}{4} + \frac{3}{4}e^t\right)^3$.

(i) Find $\mathbb{E}[X]$.

(ii) Find $\mathbb{P}(X = 2)$.

(iii) Show that X can be expressed as a sum of 3 independent observations of a random variable Y. Obtain the probability distribution of Y, and the variance of Y.

Theorem (Sum of independent normals)
If $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$ are independent, then

$$X + Y \sim N(\mu_1 + \mu_2,\ \sigma_1^2 + \sigma_2^2)$$

More generally, $aX + bY \sim N(a\mu_1 + b\mu_2,\ a^2\sigma_1^2 + b^2\sigma_2^2)$.

Example (Class practice)
$X_1, X_2, \ldots, X_n$ are independent $\text{Exp}(\lambda)$ random variables. Find the MGF of $S = X_1 + X_2 + \cdots + X_n$. (We will identify the resulting distribution in the gamma function chapter.)

The chi-squared distribution just mentioned stars in its own chapter later in the course; here, everything you need is given in the question.

Example (OCR S4, June 2018)

The random variable X has a χ^2 distribution with ν degrees of freedom. The moment generating function of X is

$$M_X(t) = (1-2t)^{-\frac{1}{2}\nu}$$

(i) Show that $\mathbb{E}[X] = \nu$.

(ii) Find $\mathrm{Var}[X]$.

(iii) Obtain the moment generating function of the sum Y of two independent χ^2 random variables, one with 6 degrees of freedom and the other with 8 degrees of freedom.

(iv) Identify the distribution of Y.

> **Remark.** The same machinery will prove the **Central Limit Theorem**: the MGF of a standardised sum of n independent copies of any (nice) random variable converges to $e^{t^2/2}$ as $n \to \infty$ – the MGF of the standard normal. See the CLT chapter.

Link with Probability Generating Functions

For a discrete random variable taking values in $\{0, 1, 2, \ldots\}$ we already have the PGF $G_X(t) = \mathbb{E}[t^X]$. Substituting $t \mapsto e^t$:

Fact —

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

so the MGF is the PGF evaluated at e^t. The MGF is the more general gadget: it makes sense for continuous variables too, where "t^X" would be unhelpful.

Example (Class practice)

The PGF of $X \sim \mathrm{Po}(\lambda)$ is $G_X(t) = e^{\lambda(t-1)}$. Verify that $G_X(e^t)$ agrees with the MGF found earlier, and check that $G_X'(1)$ and $M_X'(0)$ both give the mean.

Textbook Exercises: [S3&4] S4 Ch 4