The Sum of Two Normally Distributed Random Variables is also Normally Distributed

One of the central results in elementary statistics is that the sum (or difference) of two independent random variables each of which is normally distributed also has a normal distribution. The standard proof presented here shows that the theorem is true, but gives little insight into why it is true. It also requires quite a lot of algebra, some calculus, and almost no statistics, so students don't really learn much about the situation by going through the derivation. With those caveats, here goes:

Moment Generating Functions

First, we need to know a little about moment generating functions. A moment generating function, *M*, for a random variable *X* is defined as $M(t) = E\left(e^{tX}\right)$. This is the expected value of e^{tX}. So, for a discrete random variable, $M(t) = \sum_{all\,x} f(x)e^{tx}$ and for a continuous random variable, $M(t) = \int_{-\infty}^{\infty} f(x)e^{tx}dx$ where *f*(*x*) is the pdf for the random variable *X*. Notice that the moment generating function is a function of *t*, while the sum or integration is over *x*. This plays an important role in the manipulations that follow.

At first, this seems like a very curious definition. What possible utility could $M(t) = E\left(e^{tX}\right)$ have? The moment generating functions helps compute the moments of the distribution. Consider $M'(0)$ where differentiation is with respect to *t*. We know that

$$M'(t) = \frac{d}{dt}E\left(e^{tX}\right) = E\left(\frac{d}{dt}e^{tX}\right) = E\left(Xe^{tX}\right).$$

So, $M'(0) = E\left(Xe^{(0)X}\right) = E(X)$. The expected value, or first moment, of the distribution is the value of the first derivative of the moment generating function at $t = 0$. We can also see this by computing $M'(0) = \int_{-\infty}^{\infty} x \cdot f(x)e^{(0)x}dx$) = *E*(*X*).

The essential feature of moment generating functions is that their derivatives evaluated at $t = 0$ give you the ordinary moments of *f*, that is $M^{(n)}(0) = E\left(X^n\right)$. The expression $E\left(X^n\right)$ is called the n^{th} ordinary moment of *X*. We see that $M''(0) = E\left(X^2e^{(0)X}\right) = E\left(X^2\right)$. The variance of the distribution can be computed using the first and second ordinary moments by using the formula $E\left(X^2\right) - E(X)^2 = M''(0) - \left(M'(0)\right)^2$.

However, there is another property that makes them essential. These moments are sufficient to characterize the distributions.

Generating functions are unique! Any two random variables with the same generating functions must have the same distribution.

Moment Generating Function for the Normal Distribution

Next, we need to find the moment generating function for the normal distribution with mean μ and standard deviation σ. To find this generating function, we need to evaluate $\int_{-\infty}^{\infty} f(x)e^{tx}dx$ with $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$. So,

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} e^{tx}dx = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2+tx}dx .$$

To evaluate this integral, we need rewrite the exponent in the integrand by expanding as a quadratic in *x* and then completing the square in *x*. Considering only the exponent, we have

$$\begin{aligned}
&-\tfrac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 + tx \\
&= -\tfrac{1}{2\sigma^2}\left(x^2 - 2x\mu + \mu^2\right) + t\,x \\
&= -\tfrac{1}{2\sigma^2}\left(x^2 - 2\left(\mu - \sigma^2 t\right)x + \mu^2\right) \\
&= -\tfrac{1}{2\sigma^2}\left(\underbrace{x^2 - 2\left(\mu-\sigma^2 t\right)x + \left(\mu-\sigma^2 t\right)^2}_{factor} + \underbrace{\mu^2 - \left(\mu-\sigma^2 t\right)^2}_{simplify}\right) \\
&= -\tfrac{1}{2\sigma^2}\left(\left(x-\left(\mu-\sigma^2 t\right)\right)^2 + \left(2\mu\sigma^2 t - \sigma^4 t^2\right)\right)
\end{aligned}$$

So our integral

$$M(t) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2+tx}dx$$

becomes

$$M(t) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}\left(\left(x-\left(\mu-\sigma^2 t\right)\right)^2+\left(2\mu\sigma^2 t-\sigma^4 t^2\right)\right)}dx .$$

Since we are integrating with respect to *x*, the last two terms in the exponent are constants with respect to the integration, so pull them outside the integration. This gives

$$M(t) = e^{\left(\frac{2\mu\sigma^2 t+\sigma^4 t^2}{2\sigma^2}\right)} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\left(\mu+\sigma^2 t\right)}{\sigma}\right)^2} dx .$$

We don't really want to do the integration. And we don't have to if we recognize that the integrand is just the pdf for a normal distribution with mean $\mu + \sigma^2 t$, so the value of this integral must be 1. As a consequence, we see that the moment generating function for the normal distribution is

$$M(t) = e^{\left(\frac{2\mu\sigma^2 t + \sigma^4 t^2}{2\sigma^2}\right)} = e^{\left(\mu t + \frac{\sigma^2 t^2}{2}\right)}.$$

If $X \sim N(\mu, \sigma)$, then $M(t) = e^{\left(\mu t + \frac{\sigma^2 t^2}{2}\right)}$.

Moreover, generating functions are unique.

Any random variable whose generating function is of the form $M(t) = e^{\left(\mu t + \frac{\sigma^2 t^2}{2}\right)}$ must be normally distributed.

Before proceeding, check to see if $M'(0) = E(X) = \mu$ and $M''(0) = E(X^2) = \sigma^2 + \mu^2$.

If $M(t) = e^{\left(\mu t + \frac{\sigma^2 t^2}{2}\right)}$, then $M'(t) = e^{\left(\mu t + \frac{\sigma^2 t^2}{2}\right)}(\mu + \sigma^2 t)$. So, $M'(0) = e^{(0)}(\mu) = \mu$. So, $E(X) = \mu$. Also, using the product rule to compute the second derivative, we find

$$M''(t) = e^{\left(\mu t + \frac{\sigma^2 t^2}{2}\right)}(\sigma^2) + e^{\left(\mu t + \frac{\sigma^2 t^2}{2}\right)}(\mu + \sigma^2 t)^2.$$

So, $M''(t) = e^{(0)}\sigma^2 + e^{(0)}(\mu + 0)^2 = \sigma^2 + \mu^2$.

So, $Var(X) = E(X^2) - E(X)^2 = (\sigma^2 + \mu^2) - (\mu)^2 = \sigma^2$ as expected.

Expected Value Theorems

Now, we know that $E(X+Y) = E(X) + E(Y)$ and $E(XY) = E(X)E(Y)$ when X and Y are independent. Since the moment generating functions are expected values, $E(e^{tX})$, they obey these rules.

So, if $Z = X + Y$ (with X and Y independent and both normally distributed) then

$M_z(t) = E(e^{tZ}) = E(e^{t(X+Y)}) = E(e^{tX}e^{tY})$ by rules of exponents.

So, $E(e^{tX}e^{tY}) = E(e^{tX})E(e^{tY}) = M_X(t) \cdot M_Y(t)$, which is the product of the two moment generating functions.

Normal + Normal = Normal

So, if X and Y are independent normally distributed with means μ_x and μ_y and standard deviations σ_x and σ_y, respectively, then $Z = X + Y$ has the moment generating function

$$M_z(t) = \left(e^{\mu_x t + \frac{\sigma_x^2 t^2}{2}}\right) \cdot \left(e^{\mu_y t + \frac{\sigma_y^2 t^2}{2}}\right) = e^{(\mu_x + \mu_y)t + \frac{(\sigma_x^2 + \sigma_y^2)t^2}{2}}$$

again by rules of exponents.

But, we see that $M_z(t) = e^{(\mu_x + \mu_y)t + \frac{(\sigma_x^2 + \sigma_y^2)t^2}{2}}$ is the moment generating function for a normal distribution with mean $\mu_x + \mu_y$ and variance $\sigma_x^2 + \sigma_y^2$. Since generating functions are unique, this must mean that the sum of two normally distributed random variables is also normally distributed, which is our main theorem.

Of course, we can generalize this from a sum of two normally distributed random variables to *n* normally distributed random variables, since

$$M_z(t) = \left(e^{\mu_1 t + \frac{\sigma_1^2 t^2}{2}}\right) \cdot \left(e^{\mu_2 t + \frac{\sigma_2^2 t^2}{2}}\right) \cdot \left(e^{\mu_3 t + \frac{\sigma_3^2 t^2}{2}}\right) \cdots \left(e^{\mu_n t + \frac{\sigma_n^2 t^2}{2}}\right) = e^{(\mu_1 + \mu_2 + \mu_3 + \cdots + \mu_n)t + \frac{(\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \cdots + \sigma_n^2)t^2}{2}}.$$

So, if $X_i \sim N(\mu_i, \sigma_i)$ and $Z = X_1 + X_2 + \cdots + X_n$, then *Z* has a normal distribution with mean $\mu_1 + \mu_2 + \cdots + \mu_n$ and variance $\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2$.

Also note that if $Y \sim N(\mu, \sigma)$ then $-Y \sim N(-\mu, \sigma)$ and $M(t) = e^{\left(-\mu t + \frac{\sigma^2 t^2}{2}\right)}$. So, $Z = X - Y$ has the moment generating function

$$M_z(t) = \left(e^{\mu_x t + \frac{\sigma_x^2 t^2}{2}}\right) \cdot \left(e^{-\mu_y t + \frac{\sigma_y^2 t^2}{2}}\right) = e^{(\mu_x - \mu_y)t + \frac{(\sigma_x^2 + \sigma_y^2)t^2}{2}}$$

which is the moment generating function for a normal distribution with mean $\mu_x - \mu_y$ and variance $\sigma_x^2 + \sigma_y^2$.

Reference: Milton, J. S., and Jesse C. Arnold, *Introduction to Probability and Statistics: Principles and Applications for Engineering and Computing Science*, 2nd, McGraw-Hill, New York, 1990.