16.1 - The Distribution and Its Characteristics

16.1 - The Distribution and Its Characteristics

Normal Distribution

The continuous random variable X follows a **normal distribution** if its probability density function is defined as:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}$$

for $-\infty<x<\infty$, $-\infty<\mu<\infty$, and $0<\sigma<\infty$. The **mean** of X is μ and the **variance** of X is σ^2. We say $X\sim N(\mu,\sigma^2)$.

With a first exposure to the normal distribution, the probability density function in its own right is probably not particularly enlightening. Let's take a look at an example of a normal curve, and then follow the example with a list of the characteristics of a typical normal curve.

Example 16-1

Let X denote the IQ (as determined by the Stanford-Binet Intelligence Quotient Test) of a randomly selected American. It has long been known that X follows a normal distribution with mean 100 and standard deviation of 16. That is, $X\sim N(100,16^2)$. Draw a picture of the normal curve, that is, the distribution, of X.

https://www.youtube.com/watch/Gl7SqEaCfL0 [1]

Note that when drawing the above curve, I said "now what a standard normal curve looks like... it looks something like this." It turns out that the term "**standard normal curve**" actually has a specific meaning in the study of probability. As we'll soon see, it represents the case in which the mean μ equals 0 and the standard deviation σ equals 1. So as not to cause confusion, I wish I had said "now what a *typical* normal curve looks like...." Anyway, on to the characteristics of all normal curves!

Characteristics of a Normal Curve

It is the following known characteristics of the normal curve that directed me in drawing the curve as I did so above.

1. All normal curves are **bell-shaped** with points of inflection at $\mu\pm\sigma$.

   Proof

   The proof is left for you as an exercise

2. All normal curves are **symmetric about the mean** μ.

   Proof

All normal curves are symmetric about the mean μ, because $f(\mu+x)=f(\mu-x)$ for all x. That is:

$$f(\mu+x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{x+\mu-\mu}{\sigma}\right)^2\right\}=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right\}$$

equals:

$$f(\mu-x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{\mu-x-\mu}{\sigma}\right)^2\right\}=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{-x}{\sigma}\right)^2\right\}=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right\}$$

Therefore, by the definition of symmetry, the normal curve is symmetric about the mean μ.

3. The area under an entire normal curve is 1.

   Proof
   We prove this later on the Normal Properties page.

4. All normal curves are positive for all x. That is, $f(x)>0$ for all x.

   Proof

   The standard deviation σ is defined to be positive. The square root of 2π is positive. And, the natural exponential function is positive. When you multiply positive terms together, you, of course, get a positive number.

5. The limit of $f(x)$ as x goes to infinity is 0, and the limit of $f(x)$ as x goes to negative infinity is 0. That is:

   $\lim\limits_{x\to\infty} f(x)=0$ and $\lim\limits_{x\to-\infty} f(x)=0$

   Proof

   The function $f(x)$ depends on x only through the natural exponential function $\exp[-x^2]$, which is known to approach 0 as x approaches infinity or negative infinity.

6. The height of any normal curve is maximized at $x=\mu$.

   Proof

   Using what we know from our calculus studies, to find the point at which the maximum occurs, we must differentiate $f(x)$ with respect to x and solve for x to find the maximum. Because our $f(x)$ contains the natural exponential function, however, it is easier to take the derivative of the natural log of $f(x)$ with respect to x and solve for x to find the maximum. [The maximum of $f(x)$ is the same as the maximum of the natural log of $f(x)$, because $\log_e(x)$ is an increasing function of x. That is, $x_1<x_2$ implies that $\log_e(x_1)<\log_e(x_2)$. Therefore, $f(x_1)<f(x_2)$ implies $\log_e(f(x_1))<\log_e(f(x_2))$.] That said, taking the natural log of $f(x)$, we get:

   $$\log_e(f(x))=\log\left(\frac{1}{\sigma\sqrt{2\pi}}\right)-\frac{1}{2\sigma^2}(x-\mu)^2$$

Taking the derivative of $\log_e(f(x))$ with respect to x, we get:

$$\frac{d \log f(x)}{dx} = -\frac{1}{2\sigma^2} \cdot 2(x-\mu)$$

Now, setting the derivative of $\log_e(f(x))$ to 0:

$$\frac{d \log f(x)}{dx} = -\frac{1}{2\sigma^2} \cdot 2(x-\mu) \overset{\equiv}{\underset{SET}{}} 0$$

and solving for x, we get that $x = \mu$. Taking the second derivative of $\log_e(f(x))$ with respect to x, we get:

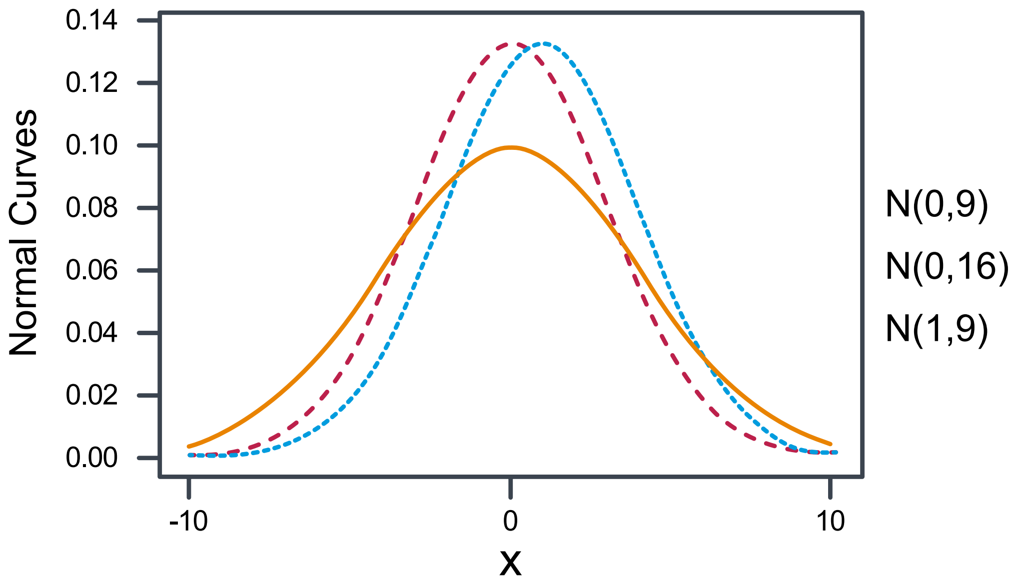
$$\frac{d^2 \log f(x)}{dx^2} = -\frac{1}{\sigma^2}$$

Because the second derivative of $\log_e(f(x))$ is negative (for all x, in fact), the point $x = \mu$ is deemed a local maximum.

7. The shape of any normal curve depends on its mean μ and standard deviation σ.

   Proof

   Given that the curve $f(x)$ depends only on x and the two parameters μ and σ, the claimed characteristic is quite obvious. An example is perhaps more interesting than the proof. Here is a picture of three superimposed normal curves —one of a $N(0,9)$ curve, one of a $N(0,16)$ curve, and one of a $N(1,9)$ curve:



   As claimed, the shapes of the three curves differ, as the means μ and standard deviations σ differ.

Legend

[1]	Link
↥	Has Tooltip/Popover
⬚	Toggleable Visibility

Source: https://online.stat.psu.edu/stat414/lesson/16/16.1

Links:

1. https://www.youtube.com/watch/Gl7SqEaCfL0