Normal Distribution and Central Limit Theorem

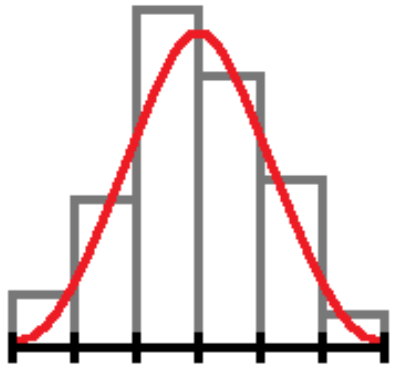
Josemari Sarasola

Statistics for Business

Normal distribution

The normal distribution is the most applied distribution in statistics:

- it's bell shaped and symmetric, and therefore it can be applied to many variables (that's because we call it *normal*);



- it's the limit of many other probability distributions;
- and it has many interesting mathematical properties.

Normal distribution

It has been studied and applied since the XVIIIth century, beginnig with the works of French mathematicians Abraham de Moivre and Pierre-Simon de Laplace. German mathematician Carl Friedrich Gauss applied it for the first time, within the research of astronomical errors. That's because is also called Gaussian distribution.

Figure: Deutsche Mark banknote: Gauss and normal distribution are depicted.

Normal distribution

Density function, parameters and notation

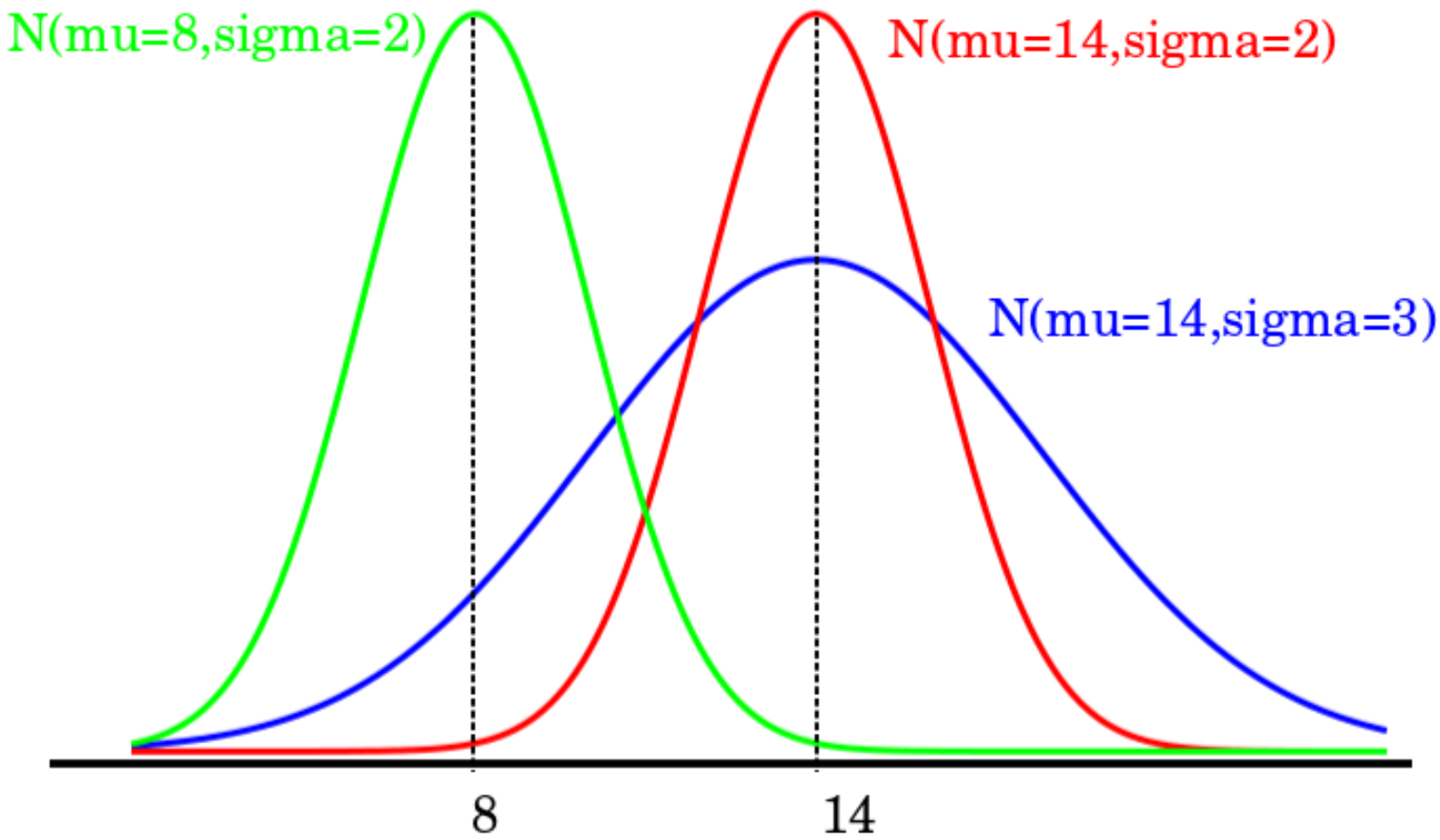
- Density function (not worth studying, it's not used):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}; \ -\infty < x < \infty$$

- $X \sim N(\mu, \sigma)$
  μ : expected value σ : standard deviation

It's symmetrical around μ (very useful to calculate probabilities)

Normal distribution



Normal distribution

Standard normal distributions

$Z \sim N(\mu = 0, \sigma = 1)$ is the *standard* normal distribution. We always call it Z. We use it as a basis to calculate probabilities for any other normal distribution, by means of *standardizing* (see next slide).

Standardizing

To calculate probabilities easily, all normal distribution must be transformed to the standard normal distribution, by menas of *standardizing*:

$$X \sim N(\mu, \sigma) \rightarrow Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

So, if we take any normal distribution and we substract its mean and divide it by its standard deviation, it will become a standard normal distribution. We call the standardized values standard scores.

Linear transformations

Let be $X \sim N(\mu, \sigma)$ and let's create this transformed variable: $Y = a + bX$. It can be proven that Y is also normal:

$$Y \sim N(a + b\mu, |b|\sigma)$$

Sum of normal distributions

Let be

- $X_1 \sim N(\mu_1, \sigma_1)$
- $X_2 \sim N(\mu_2, \sigma_2)$
- ...
- $X_n \sim N(\mu_n, \sigma_n)$ all of them independent with each other.

It can be proven:

$$Y = X_1 + X_2 + ... + X_n \sim N\left(\mu_1 + \mu_2 + ... + \mu_n, \sigma = \sqrt{\sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2}\right)$$

Hence, sum of normal distributions follows a normal distribution. In statistics we say the normal distribution is reproductive, and the property is named reproductivity.

Remark: $\sqrt{\sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2} \neq \sigma_1 + \sigma_2 + ... + \sigma_n$ (we always add variances!)

Normal distribution

R software

- `pnorm(1.45)` `#P[Z<1.45]`
- `1-pnorm(1.45)` `#P[X>1.45]`
- `pnorm(1.45,lower.tail=FALSE]` `#P[X>1.45]`
- `pnorm(4,mean=3,sd=1.5)` `#P[X>4], X:N(3,1.5)`
- `qnorm(0.90)` `#P[Z<z]=0.9,z?`
- `qnorm(0.90,,mean=3,sd=1.5)` `#P[X<x]=0.9, x?,X: N(3,1.5)`

De Moivre-Laplace theorem:
normal approximation of the binomial distribution

If we have a $B(n,p)$ distribution, with big n $(n \geq 30)$, and p not being small $(np \geq 10)$, -because in those cases, with small p, we use Poisson approximation, taking $\lambda = np$-, we can approximate binomial probabilities with a normal distribution:

$$B(n,p) \xrightarrow[n \geq 30]{} N(\mu = np, \sigma = \sqrt{npq})$$

This approximation is called De Moivre-Laplace theorem.

Normal approximation of the Poisson distribution

If we have a $P(\lambda)$ distribution, with a big λ ($\lambda \geq 30$), we can approximate it with a normal distribution:

$$P(\lambda) \xrightarrow[\lambda \geq 30]{} N(\mu = \lambda, \sigma = \sqrt{\lambda})$$

Continuity correction

Applying De Moivre-Laplace and normal approximation for Poisson, we are finally approximating a discrete distribution (binomial and Poisson) by a continuous distribution (normal). In order to be more precise in the calculations, we use the continuity correction. Here you have some examples:

- $P[X = 10] = P[9.5 < X < 10.5]$
- $P[X \leq 10] = P[X < 10.5]$
- $P[X \geq 10] = P[X > 9.5]$
- $P[X > 10] = P[X > 10.5]$
- $P[X < 10] = P[X < 9.5]$

The error resulting from not applying the corrections is most times very small. Therefore, the correction is needed only when we have to be very precise in the calculations.

Central Limit Theorem (CLT)

We know that sum of normal distributions distributes following a normal distribution. But what if adding up distributions are not normal?

In that case, sum of distributions distributes normally too, but if two condtions are held:

1. distributions must be independent (as for sum of normal distributions) and,
2. no. of adding up distributions must be large (generally, 30 or larger).

We name this result Central Limit Theorem (CLT). De Moivre-Laplace theorem and normal approximation for Poisson are special cases of CLT.

Central Limit Theorem (CLT)

- $X_1 \sim ?(\mu_1, \sigma_1)$
- $X_2 \sim ?(\mu_2, \sigma_2)$
- ...
- $X_n \sim ?(\mu_n, \sigma_n)$, all of them independent, with known means and variances, and having $n \geq 30$,
- then, this holds:

$$Y = X_1 + X_2 + ... + X_n \sim N\left(\mu_1 + \mu_2 + ... + \mu_n, \sigma = \sqrt{\sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2}\right)$$

That is to say, sum of many independent distributions is always a normal distribution, taken sum of means as the mean, and sum of variances as the variance.