Statistics

1 Linear combinations of random variables

Continuous random variables

A continuous random variable X has a pdf f such that:

1. $f(x) \geq 0 \forall x$
2. $\int_{-\infty}^{\infty} f(x)\,dx = 1$

$$\Pr(X \leq c) = \int_{-\infty}^{c} f(x)\,dx$$

Linear functions $X \to aX + b$

$$\begin{aligned}\Pr(Y \leq y) &= \Pr(aX + b \leq y) \\ &= \Pr\left(X \leq \frac{y-b}{a}\right) \\ &= \int_{-\infty}^{\frac{y-b}{a}} f(x)\,dx\end{aligned}$$

Mean:	$\mathrm{E}(aX + b) = a\,\mathrm{E}(X) + b$
Variance:	$\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$

Linear combination of two random variables

Mean:	$\mathrm{E}(aX + bY) = a\,\mathrm{E}(X) + b\,\mathrm{E}(Y)$	
Variance:	$\mathrm{Var}(aX + bY) = a^2\,\mathrm{Var}(X) + b^2\,\mathrm{Var}(Y)$	(if X and Y are independent)

2 Sample mean

Approximation of the **population mean** determined experimentally.

$$\overline{x} = \frac{\Sigma x}{n}$$

where n is the size of the sample (number of sample points) and x is the value of a sample point

On CAS

1. Spreadsheet
2. In cell A1: `mean(randNorm(sd, mean, sample size))`
3. Edit → Fill → Fill Range
4. Input range as A1:An where n is the number of samples
5. Graph → Histogram

Sample size of n

$$\overline{X} = \sum_{i=1}^{n} \frac{x_i}{n} = \frac{\sum x}{n}$$

Sample mean is distributed with mean μ and sd $\frac{\sigma}{\sqrt{n}}$ (approaches these values for increasing sample size n).

On CAS

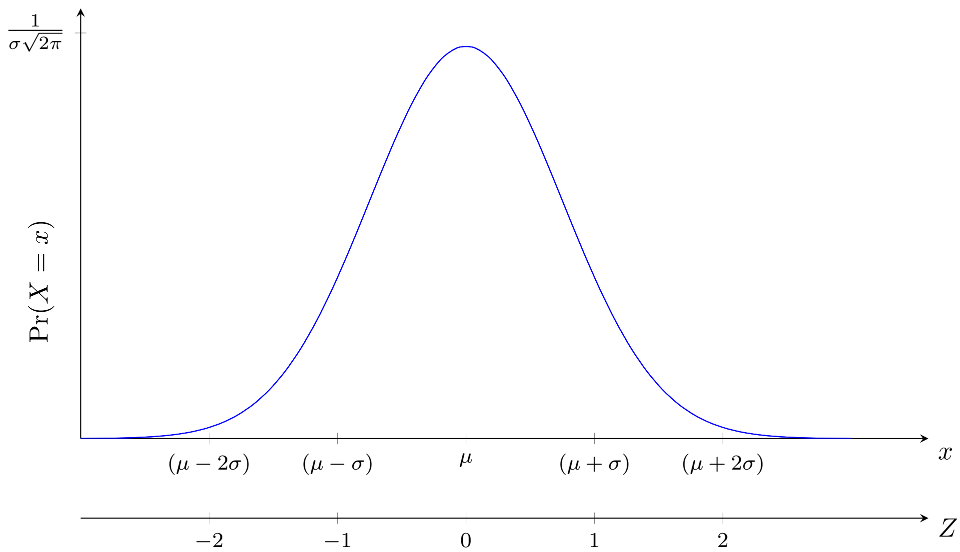
- Spreadsheet → Catalog → `randNorm(sd, mean, n)` where `n` is the number of samples. Show histogram with Histogram key in top left
- To calculate parameters of a dataset: Calc → One-variable

3 Normal distributions

mean = mode = median

$$Z = \frac{X - \mu}{\sigma}$$

Normal distributions must have area (total prob.) of 1 $\implies \int_{-\infty}^{\infty} f(x)\, dx = 1$



4 Central limit theorem

If X is randomly distributed with mean μ and sd σ, then with an adequate sample size n the distribution of the sample mean $\overline{X}$ is approximately normal with mean $E(\overline{X})$ and $\text{sd}(\overline{X}) = \frac{\sigma}{\sqrt{n}}$.