Uni-variate (1D) random variables (complete description)

Random variable X

Outcome (value) $X = x$

Probability density function (pdf) $f_X(x) = \lim_{dx \to 0} \frac{P(X \in [x, x+dx])}{dx}$

$\int_{-\infty}^{\infty} f_X(x) dx = 1\,,\; f_X(x) \geq 0$

$P(X \in [a,b]) = \int_a^b f_X(x) dx$

Cumulative density function (cdf) $F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(x) dx$

$F_X(-\infty) = 0,\;\; F_X(\infty) = 1$

$P(X \in [a,b]) = F_X(b) - F_X(a)$

Domain and range: all random variables can be considered continuous with possible values $x \in R$, since discrete random variables can be modeled with impulse-functions in their pdf (and step-functions in their cdf) and since the pdf can equal zero in "impossible" x-regions

---

Population, i.e. theoretical, moments (summarizing, incomplete description)

Population mean (1'st order) $\mu_X = E[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)\, dx$

measures location (of probability mass)

Population variance (2'nd order) $\sigma_X^2 = E[(X - \mu_X)^2] = \int_{-\infty}^{\infty} (x - \mu_X)^2 \cdot f_X(x)\, dx$

$\sigma_X^2 = E[X^2] - \mu_X^2$

measures dispersion (of probability mass)

Population standard deviation $\sigma_X = \sqrt{\sigma_X^2}$

Descriptive statistics and sample moments

Sample (observed data values) $X = \{x_1, x_2, \dots x_n\}$

Sample size n

Sample mean (average) $\hat{\mu}_X = \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$

Sample variance $\hat{\sigma}_X^2 = S_X^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$

Sample standard deviation $\hat{\sigma}_X = \sqrt{\hat{\sigma}_X^2}$