Lectures

MVE055 / MSG810
Mathematical statistics and discrete mathematics

Moritz Schauer

Last updated September 6, 2022

GU & Chalmers University of Technology

Continuous distributions

Continuous random variables

Continuous random variables

A continuous random variable can assume all values in one or several intervals of real numbers, and the probability of assuming a particular value is zero.

Examples:

```
julia> rand()
0.8891381949344658
```

Continuous distributions

A continuous random variable X is described by its *probability density function (pdf)* $f(x)$

$$\mathsf{P}(a \leq X \leq b) = \int_a^b f(x) \mathrm{d}x.$$

$$\mathsf{P}(X = x) = 0$$

and

$$\mathsf{P}(a \leq X \leq b) = \mathsf{P}(a < X \leq b) = \mathsf{P}(a \leq X < b) = \mathsf{P}(a < X < b)$$

Continuous distributions

Probability density function (pdf)

A function is a probability density function (pdf) if and only if

$$f(x) \geq 0 \quad \text{and} \quad \int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1.$$

Example

Show that the function

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

is a pdf.

$$f(x) \geq 0 \quad \checkmark.$$

$$\begin{aligned} \int_{-\infty}^{+\infty} f(t)dt &= \int_{-\infty}^{a} 0 \, \mathrm{d}t + \int_{a}^{b} \frac{1}{b-a} \, \mathrm{d}t + \int_{b}^{\infty} 0 \, \mathrm{d}t \\ &= \int_{a}^{b} \frac{1}{b-a} \mathrm{d}t = \frac{b-a}{b-a} = 1 \quad \checkmark. \end{aligned}$$

Continuous distributions

Cumulative distribution function

The cumulative distribution function F of a continuous distribution is

$$F(x) = \mathsf{P}(X \leq x) = \int_{-\infty}^{x} f(t)\mathrm{d}t.$$

$$\mathsf{P}(a \leq X \leq b) = F(b) - F(a)$$

Example

Find cumulative distribution function for X with pdf

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t = \begin{cases} 0 & x \leq a \\ \frac{x-a}{b-a} & x \in [a,b] \\ 1 & x \geq b. \end{cases}$$

Expected value

The expected value is an "average" outcome of a random variable.

Expected value

The expected value of a random variable is defined as

$$\mathsf{E}[X] = \begin{cases} \int\limits_{-\infty}^{\infty} x f(x) \mathrm{d}x & \text{if } X \text{ is continuous,} \\ \sum\limits_{\text{all } x} x f(x) & \text{if } X \text{ is discrete.} \end{cases}$$

Rules for computing expected values

For the expected value,

- $\mathsf{E}[a) = a$.
- $\mathsf{E}[aX] = a\mathsf{E}[X]$.
- $\mathsf{E}[aX + b] = a\mathsf{E}[X] + b$.
- $\mathsf{E}[X + Y] = \mathsf{E}[X] + \mathsf{E}[Y]$.

Here X and Y are two random variables and a and b are constants.

The *same* rules: E is a linear operator on random variables.

Uniform distribution

Uniform distribution

The continuous distribution with pdf
$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases} .$$

is called the *uniform distribution*. Facts: $\mathsf{E}[X] = (a+b)/2$.

$$\mathsf{E}[X] = \int\limits_{-\infty}^{\infty} x f(x) \mathrm{d}x = \frac{1}{b-a} \int_a^b x \mathrm{d}x = \frac{\frac{1}{2}b^2 - \frac{1}{2}a^2}{a-b} = (a+b)/2$$

Transformations

If we transform the random variables by a function h we have:

Theorem ♡

$$\mathsf{E}[h(X)] = \begin{cases} \sum\limits_{\text{all } x} h(x)f(x), & \text{if } X \text{ is discrete,} \\ \cdots & \\ \int\limits_{-\infty}^{\infty} h(x)f(x)\mathrm{d}x, & \text{if } X \text{ is continuous.} \end{cases}$$

Variance

Variance and standard deviation

Variance

The variance of a random variable is defined as

$$\mathsf{V}(X) = \mathsf{E}[(X-\mu)^2],$$

where $\mu = \mathsf{E}[X]$ is the expected value of X.

In words, this is the expected squared deviation of the mean. The variance can be calculated by

$$\mathsf{V}(X) = \begin{cases} \sum_{\text{all } k} (k-\mu)^2 f(k), & \text{for discrete } X \\ \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\mathrm{d}x, & \text{for continuous } X. \end{cases}$$

Sometimes it is easiest to compute $\mathsf{V}(X) = \mathsf{E}[X^2] - \mu^2$.

The standard deviation of a random variable X is defined as $\sigma = \sqrt{\mathsf{V}(X)}$. It has the same unit as X.

Rules for computing variance

For the variance

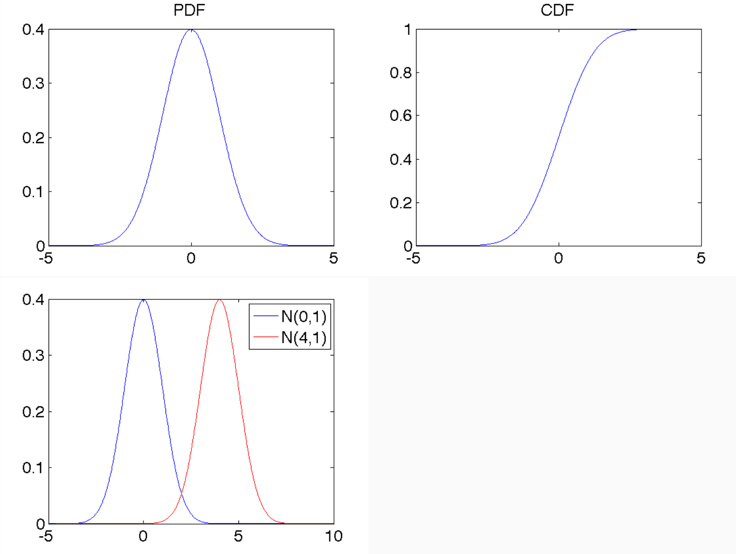
- $\mathsf{V}(a) = 0$.
- $\mathsf{V}(aX) = a^2\mathsf{V}(X)$.
- $\mathsf{V}(aX + b) = a^2\mathsf{V}(X)$.
- $\mathsf{V}(X + Y) = \mathsf{V}(X) + \mathsf{V}(Y)$, if X and Y are independent.

Here X and Y are two random variables and a and b are constants.

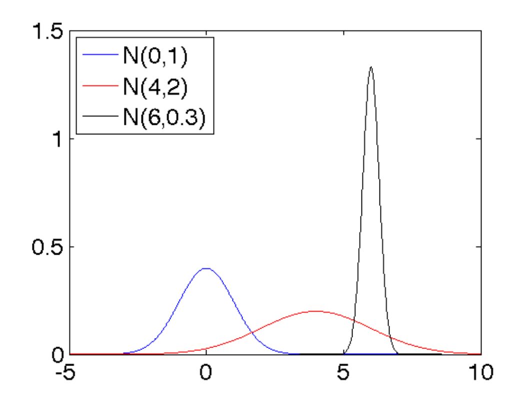
Normal distributions

Normal distribution

Density and distribution function of $Z \sim \mathrm{N}(0,1)$ and $\mathrm{N}(4,1)$



pdf's for some other possible parameters



Normal distribution

Normal distribution $\mathsf{N}(\mu, \sigma^2)$

A r.v. X is normally distributed, $X \sim \mathsf{N}(\mu, \sigma^2)$, with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$, if it has pdf

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

The distribution function is

$$F(x) = \int_{-\infty}^{x} = \ldots \text{has no nice solution}$$

Parameters

If $X \sim \mathsf{N}(\mu, \sigma^2)$ then $\mathsf{E}[X] = \mu$ and $\mathsf{V}(X) = \sigma^2$.

Normal distribution pdf

Standard normal distribution

Standard normal distribution

A r.v. random variable Z is standard normally distributed if $Z \sim \mathsf{N}(0, 1)$. Then $\mathsf{E}[Z] = 0$ and $\mathrm{Var}(Z) = 1^2$.

We denote pdf and cdf by $\varphi(x)$ and $\Phi(x)$

Normalisation

Theorem

If $X \sim \mathsf{N}(\mu, \sigma^2)$ then $aX + b \sim \mathsf{N}(a\mu + b, a^2\sigma^2)$.

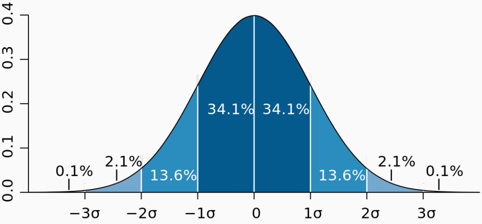
That means for $X \sim \mathsf{N}(\mu, \sigma^2)$ that

- $X = \mu + \sigma Z$ where $Z \sim \mathsf{N}(0, 1)$.
- $Z = (X - \mu)/\sigma \sim \mathsf{N}(0, 1)$.

We use this to sample random variables, and to compute probabilities:

$$\mathsf{P}(X < x) = \mathsf{P}\left(\frac{X - \mu}{\sigma} < \frac{x - \mu}{\sigma}\right) = \mathsf{P}\left(Z < \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right).$$

Rule



Example: IQ values are normalized such that (approximately)

$$\mathrm{IQ} \sim \mathsf{N}(100, 15^2)$$

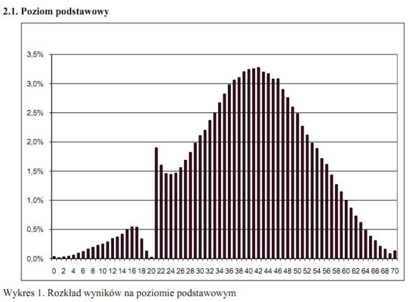
What is the probability that a random person scores 115 or more?
Approx. $13.6 + 2.1 + 0.1 = 15.8$.

Relict of the past: Normal distribution table

Table gives $\Phi(z) = \mathsf{P}(X \leq z)$ for $Z \sim \mathsf{N}(0,1)$.
For negative values use that $\Phi(-z) = 1 - \Phi(z)$.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0 :	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1 :	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2 :	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3 :	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4 :	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5 :	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6 :	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7 :	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8 :	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9 :	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0 :	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1 :	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2 :	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3 :	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4 :	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5 :	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6 :	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7 :	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8 :	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9 :	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0 :	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1 :	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2 :	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3 :	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4 :	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5 :	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952

High-school maturity exam in Poland



Histogram showing the distribution of scores for the obligatory Polish language test. "The dip and spike that occurs at around 21 points just happens to coincide with the cut-off score for passing the exam"

`http://freakonomics.com/2011/07/07/`

`another-case-of-teacher-cheating-or-is-it-just-altruism/`

Moment generating function (m.g.f.)

Let X be a random variable

- The k^{th} moment for X is defined by $\mathsf{E}[X^k]$.
- The moment generating function for X is defined by

$$m_X(t) = \mathsf{E}[\mathrm{e}^{tX}].$$

- Let $m_X(t)$ be the m.g.f for X. Then

$$\left.\frac{\mathrm{d}^k m_X(t)}{\mathrm{d}t^k}\right|_{t=0} = \mathsf{E}\left[X^k\right]$$

Moment generating function for standard normal distribution

Let $Z \sim \mathsf{N}(0,1)$. Compute the mgf. Use $h(x) = \mathrm{e}^{tx}$ and transform:

$$m_X(t) = \mathsf{E}\left[e^{tX}\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \underbrace{\mathrm{e}^{tx}}_{h(x)} \mathrm{e}^{-\frac{1}{2}x^2} \mathrm{d}x$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-\frac{1}{2}(x-t)^2} \mathrm{e}^{\frac{1}{2}t^2} \mathrm{d}x = \mathrm{e}^{\frac{1}{2}t^2}$$