Matematisk Statistik och Disktret Matematik, MVE055/MSG810, HT19

Föreläsning 3

Nancy Abdallah

Chalmers - Göteborgs Universitet

September 9, 2019

Continuous Random Variables

- A **continuous random variable** can take all the values in an interval of real numbers (or all real values).
- If X is a continuous random variable $P(X = x) = 0$ and $P(a \leq x \leq b) \geq 0$, for all $a, b \in \mathbb{R}$.
- For every continuous random variable there exists a function $f(x)$ such that

$$P(a \leq x \leq b) = \int_a^b f(x)dx$$

  $f(x)$ is called a **density function** *(sv. täthetsfunktion)*.
- If X is a continuous random variable with density function *f(x)*, then

$$P(a \leq x \leq b) = P(a < x \leq b) = P(a \leq x < b) = P(a < x < b)$$

- A function $f(x)$ is a density function for a continuous random variable if and only if
  - (i) $f(x) \geq 0$ for all $x \in \mathbb{R}$, and
  - (ii) $\int_{-\infty}^{+\infty} f(x) = 1$
- The cumulative distribution function $F(x)$ is defined by

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

The distribution function of a continuous random variable is continuous.

- At every point x where $f(x)$ is continuous,

$$F'(x) = f(x)$$

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 density function.
Solution:
$f(x) \geq 0$ *for all* $x \in \mathbb{R}$ *and*

$$\int_{-\infty}^{+\infty} f(t)dt = \int_a^b \frac{1}{b-a}dt = \frac{b-a}{b-a} = 1$$

Therefore $f(x)$ *is a density function. The random variable whose density function is given above is said to have a* ***uniform distribution*** *(sv. likformig fördelning).*

Example

Let X be a continuous random variable with density function

$$f(x) = \begin{cases} 12.5x - 1.25 & if\, 0.1 \le x \le 0.5 \\ 0 & otherwise \end{cases}$$

Find the cumulative distribution function for X and compute $P(0.3 \le X \le 0.6)$.

Solution:

$$F(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 0 & om\ x < 0.1 \\ \int_{0.1}^{x}(12.5t - 1.25)dt & om\, 0.1 \le x < 0.5 \\ 1 & om\ x > 0.5 \end{cases}$$

Example

Therefore,

$$F(x) = \begin{cases} 0 & om\ x < 0.1 \\ 6.25x^2 - 1.25x + 0.0625 & om\ 0.1 \leq x < 0.5 \\ 1 & om\ x > 0.5 \end{cases}$$

We can use $F(x)$ to compute $P(a \leq X \leq b)$.

$P(0.3 \leq X \leq 0.6) = F(0.6) - P(0.3) = 1 - 0.25 = 0.75.$

Or, without using $F(x)$,

$$\begin{aligned} P(0.3 \leq X \leq 0.6) &= \int_{0.3}^{0.6} f(x)dx = \int_{0.3}^{0.5} (12.5x - 1.25)dx \\ &= [12.5\frac{x^2}{2} - 1.25x]_{0.3}^{0.5} = 0.75 \end{aligned}$$

Expected value, Variance, Standard deviation

Let X be a continuous random variable with density function $f(x)$.

- The expected value of X is given by

$$E[X] = \int_{-\infty}^{+\infty} xf(x)dx.$$

- In general, if $H(X)$ is a random variable, then the expected value of $H(X)$ is given by

$$E[H(X)] = \int_{-\infty}^{+\infty} H(x)f(x)dx$$

- The variance and the standard deviation are defined in the same way as for discrete random variables, i.e. $Var[x] = E[X^2] - E[X]^2$, and $\sigma = \sqrt{Var[x]}$.

Example

Let X be a continuous random variable with density function

$$f(x) = \begin{cases} 12.5x - 1.25 & \text{if } 0.1 \le x \le 0.5 \\ 0 & \text{otherwise} \end{cases}$$

The expected value for X is

$$\begin{aligned} \mu = E[X] &= \int_{-\infty}^{\infty} f(x)dx = \int_{0.1}^{0.5} x(12.5x - 1.25) \\ &= \left[\frac{12.5x^3}{3} - \frac{1.25x^2}{2} \right]_{0.1}^{0.5} \\ &= 0.3667 \end{aligned}$$

The rules for the expected value and the variance of a continuous random variable are the same as those for a discrete random variable. That is, for two random variables X and Y and a constant c,

- $E[c] = c$
- $E[cX] = cE[X]$
- $E[X + Y] = E[X] + E[Y]$
- $Var[c] = 0$
- $Var[cX] = c^2 Var[X]$
- If X and Y are independent then $Var[X + Y] = Var[X] + Var[Y]$

Normal distribution

- A random variable with density function

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

  where $\sigma > 0$ and $x, \mu \in \mathbb{R}$, i said to have a normal distribution with parameters μ and σ.
- Notation: $X \sim N(\mu, \sigma^2)$.
- $E[X] = \mu$ and $Var[X] = \sigma^2$.

Graph of normally distributed random variables

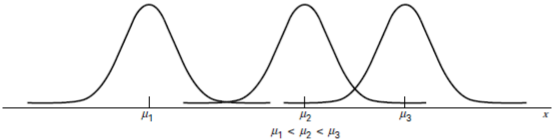


FIGURE 4.6.3 Three normal distributions with different means but the same amount of variability.

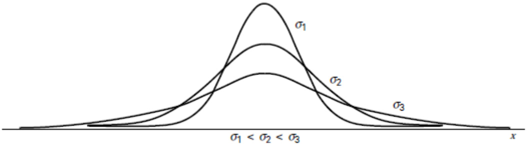
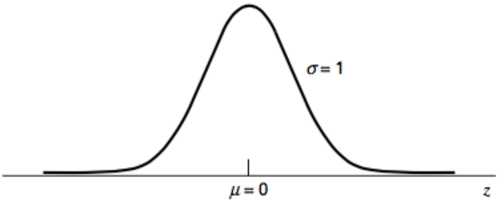


FIGURE 4.6.4 Three normal distributions with different standard deviations but the same mean.

Standard Normal distribution

If $X \sim N(0, 1)$, X is said to have a standard normal distribution, and its density function $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ is given in the graph below.

Remark: A standard normal distribution is usually denoted by Z instead of X and it's graph is symmetric with respect to the vertical line $z = \mu$.



Let $Z \sim N(0,1)$. To compute $P(a < Z < b)$ where a and b are two real numbers (that can be infinite), we use Table V s.697-698 of the cumulative distribution function $F(x)$.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
$\vdots$										
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015

Examples using the table above:

$P(Z \leq 1.24) = 0.8925$

$P(Z > 1.2) = 1 - P(Z \leq 1.2) = 1 - 0.8849 = 0.1151$

$P(Z \leq -1.2) = P(Z \geq 1.2) = P(Z > 1.2) = 0.1151$

Theorem

Suppose X is normally distributed with mean μ and standard deviation σ. The variable $\frac{X-\mu}{\sigma}$ is standard normal.

Example

Let $X \in N(17, 5)$ and suppose we want to find $P(X \leq 20)$. Let $Z = (X-17)/\sqrt{5}$. Z is standard normal.

$$\begin{aligned} P(X \leq 20) &= P\left(\frac{X-17}{\sqrt{5}} \leq \frac{20-17}{\sqrt{5}}\right) = P(Z \leq 1.34) \\ &= F(1.34) = 0.9099 \end{aligned}$$

For which value of x is $P(X > x) = 0.6$?

$P\left(\frac{X-17}{\sqrt{5}} > \frac{x-17}{\sqrt{5}}\right) = P\left(z > \frac{x-17}{\sqrt{5}}\right) = 0.6$

$\Rightarrow P\left(z < \frac{x-17}{\sqrt{5}}\right) = 1 - 0.6 = 0.4 \Rightarrow \frac{x-17}{\sqrt{5}} \approx -0.255$

Hence $x \approx -0.255\sqrt{5} + 17 = 16.43$

Normal approximation to the binomial distribution

Theorem

Let $X \in Bin(n, p)$. If [$p \le 0.5$ and $np > 5$] or [$p > 0.5$ and $n(1-p) > 5$], then X is approximately normally distributed with mean np and variance $np(1-p)$).

Remark

Notice that a binomial distribution is discrete and a normal distribution is continuous. Therefore, for more precision

$$P(X \le x) \approx P(Y \le x + \frac{1}{2})$$

and

$$P(X < x) \approx P(Y \le x - \frac{1}{2}).$$

Transformation av kontinuerliga s.v.

- Suppose that X is a continuous random variable with density function f_X and assume that the variable Y is defined such that $h(Y) = X$ where h is strictly monotonic and differentiable function. Then

$$f_Y(y) = f_X(h(y))|h'(y)|$$

- Example: If $X = aY + b$ then $f_Y(y) = f_X(ay + b)|a|$.
- Example: If $X \in N(\mu, \sigma^2)$ and $Y = \frac{X-\mu}{\sigma}$ then $X = \sigma Y + \mu$ and

$$\begin{aligned} f_Y(y) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}((\sigma y + \mu) - \mu)^2\right)\sigma \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) \end{aligned}$$