

Biomeetria

Kahe arvtunnuse ühine käitumine, korrelatsioonanalüüs

Lineaarne e Pearsoni korrelatsioonikordaja



Millal kasutada ja mida näitab?


Mõõdab kahe pideva (normaaljaotusega) tunnuse vahelise linearse seose tugevust ja suunda.

Arvutusvalem:
$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}}$$

Omadused:

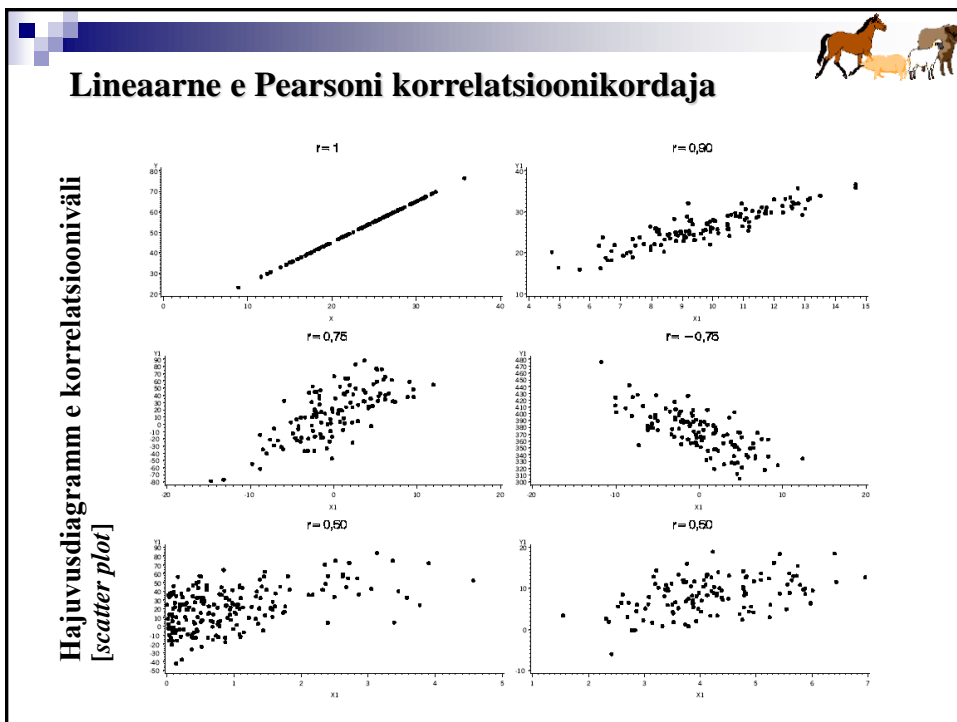
- $-1 \leq r \leq 1$;
- kui $r > 0$, siis tunnuse X suurenedes suureneb keskmiselt ka tunnus Y ;
kui $r < 0$, siis X -i suurenedes Y keskmiselt kahaneb ja X -i kahanedes Y keskmiselt suureneb;
- kui tunnused X ja Y on sõltumatud, siis $r = 0$;
- kui tunnuste X ja Y vahel on täpne lineaarne seos, siis $|r| = 1$;
- mida suurem on korrelatsioonikordaja absoluutväärtus, seda tugevam on korrelatiivne seos tunnuste vahel.

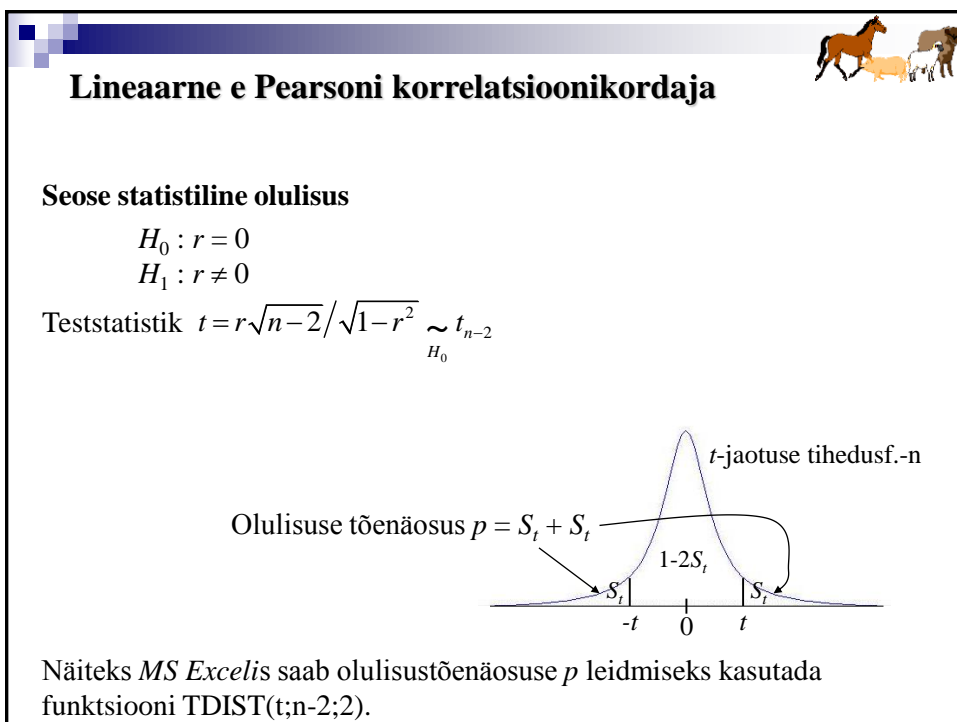
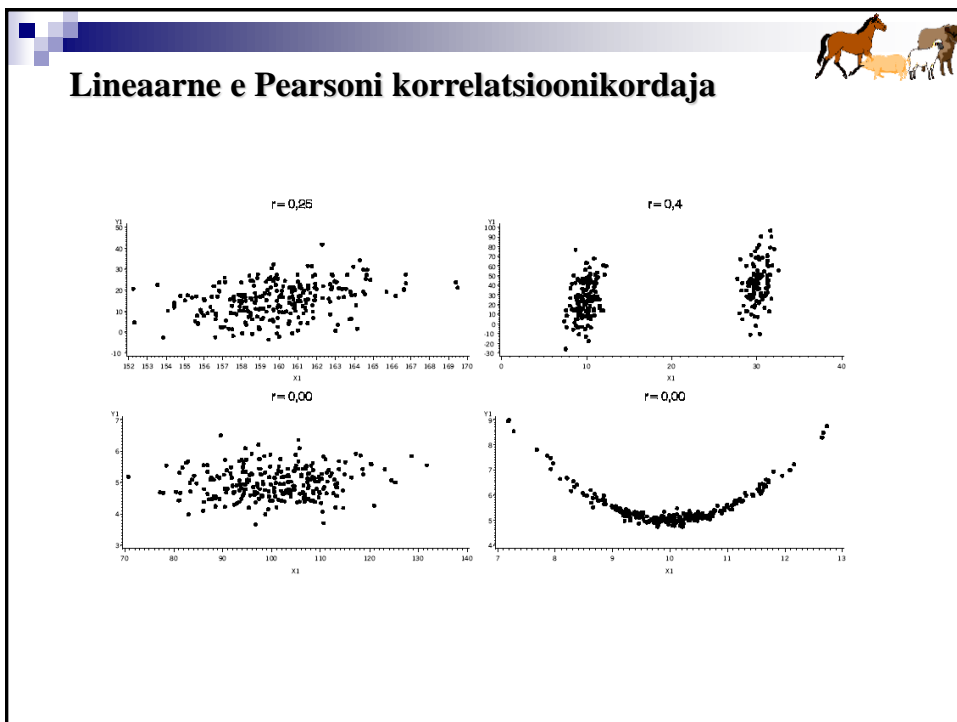
Lineaarne e Pearsoni korrelatsioonikordaja

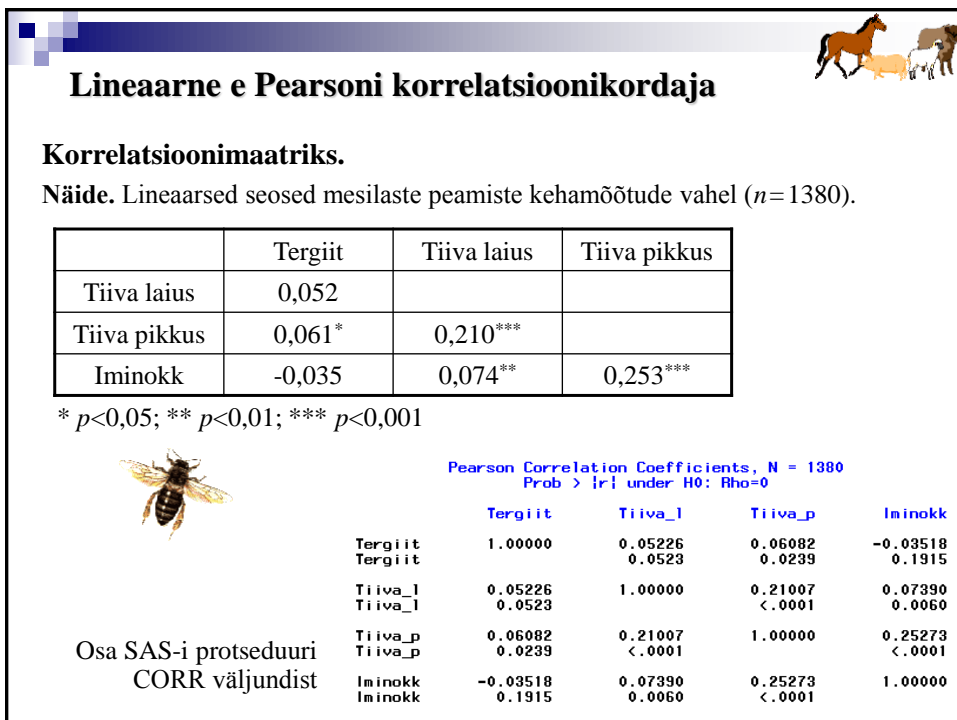
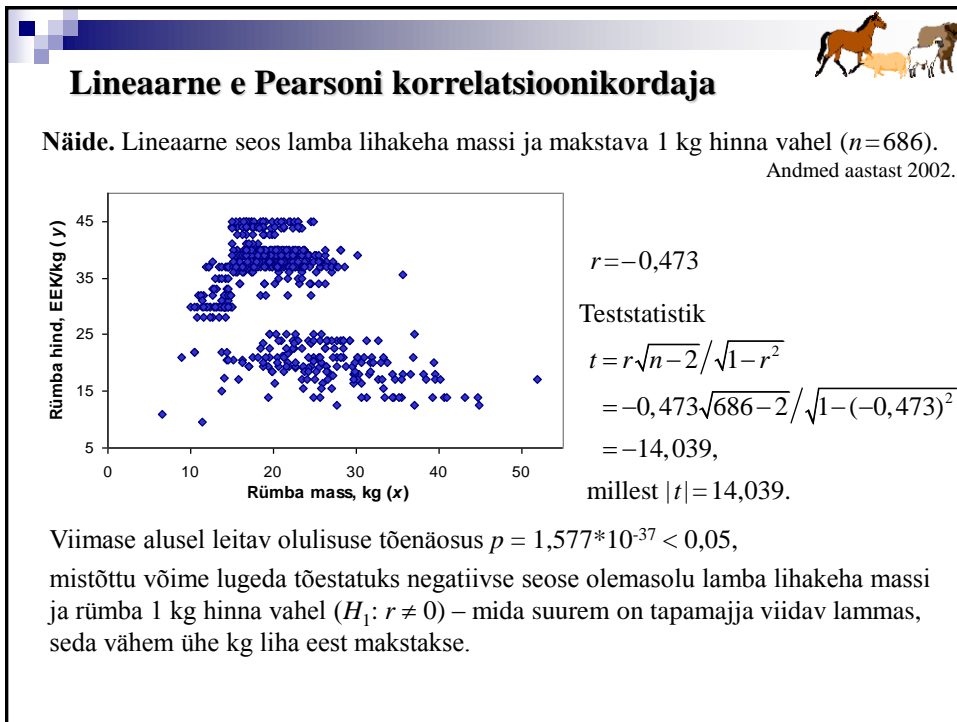


Kokkuleppelised piirid seose tugevuse iseloomustamiseks:

- $|r| < 0,3$ – nõrk seos;
- $0,3 \leq |r| < 0,7$ – keskmine seos;
- $|r| \geq 0,7$ – tugev seos.







Astakkorrelatsioonikordaja e Spearmani korrelatsioonikordaja



Millal kasutada ja mida näitab?

Mõõdab kahe arvunnuse vahelise monotoonse seose tugevust ja suunda.

Ei ole tundlik erindite suhtes ega eelda tunnuste normaalset jaotumist (on põhimõtteliselt kasutatav ka järjestustunnuste puhul).

Arvutamine: leitakse kui seos vaatluste järjekorranumbrite e astakute vahel:

$$\rho = 1 - \frac{6 \sum_{i=1}^n (x_{(i)} - y_{(i)})^2}{n(n^2 - 1)}, \quad \text{kus } x_{(i)} \text{ on tunnuse } X \text{ väärtuse } x_i \text{ astak} \\ \text{ja } y_{(i)} \text{ on tunnuse } Y \text{ väärtuse } y_i \text{ astak.}$$

Omadused:

- $-1 \leq \rho \leq 1$;
- kui tunnuste vahel on kasvav monotoonne seos, siis $\rho > 0$;
kui tunnuste vahel on kahanev monotoonne seos, siis $\rho < 0$;
- kui tunnused X ja Y on sõltumatud, siis $\rho = 0$;
- kui tunnuste X ja Y vahel on funktsionaalne monotoonne seos, siis $|\rho| = 1$.

Lineaarne versus astakkorrelatsioonikordaja

