

Lineare Regression

Dokumentnummer: DX1233
 Fachgebiet: beschreibende Statistik
 Einsatz: 2HAK (erstes Lernjahr)

```
(%i1) x:[10,20,40,70,200];
      y:[21,41,81,141,401];
(%o1) [10, 20, 40, 70, 200]
(%o2) [21, 41, 81, 141, 401]
```

```
(%i3) n:length(x);
(%o3) 5
```

```
(%i4) mx:sum(x[i],i,1,n)/n;
      my:sum(y[i],i,1,n)/n;
(%o4) 68
(%o5) 137
```

Figure 1: Berechnung der Kovarianz

Covarianz:

$$\text{Cov}(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \cdot \bar{y}$$

```
(%i6) cov_x_y:1/n*sum((x[i]-mx)*(y[i]-my),i,1,n);
(%o6) 9552
```

```
(%i7) cov_x_y:1/n*sum(x[i]*y[i]-mx*my,i,1,n);
(%o7) 9552
```

Figure 2: Koeffizienten der Regressionsgeraden bestimmen

Regressionsgerade:
 $y = ax + b$

$$a = \frac{\text{Cov}(x, y)}{V(x)}$$

$$b = \bar{y} - a\bar{x}$$

```
(%i11) V_x:sum((x[i]-mx)**2,i,1,n)/n;
(%o11) 4776
```

```
(%i12) a:cov_x_y/V_x;
(%o12) 2
```

```
(%i10) b:my-a*mx;
(%o10) 1
```

Figure 3: Korrelationskoeffizient bestimmen

Korrelationskoeffizient:

$$r = \frac{\text{Cov}(x,y)}{\sqrt{V(x) \cdot V(y)}}$$

```
(%i13) V_y:sum((y[i]-my)**2,i,1,n)/n;
```

```
(%o13) 19104
```

```
(%i14) r:cov_x_y/sqrt(V_x*V_y);
```

```
(%o14) 1
```