Stichprobenvarianz

Donnerstag, 28. Oktober 2021 16:41

Aufgabe 29 *Stichprobenvarianz* **5 P.**

Für alle Berechnungen sind $x_1, \dots, x_n$ die Ausprägungen der quadratisch integrierbaren, paarweise unkorrelierten, reellwertigen Zufallsvariablen $X_1, \dots, X_n$ mit der Varianz σ^2 und dem Mittelwert μ.

(a) Testen Sie, ob die Formel (arithmetisches Mittel)

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{1}$$

eine erwartungstreue Schätzfunktion für den Mittelwert μ der Grundgesamtheit ist. Falls die Schätzfunktion nicht erwartungstreu ist, suchen Sie nach einer geeigneten Korrektur.

(b) Der Standardfehler des arithmetischen Mittels (1) ist definiert als die Wurzel aus der Varianz von $\bar{X}$. Zeigen Sie, dass

$$\mathrm{E}\left((\bar{X}-\mu)^2\right) = \mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n} \tag{2}$$

gilt.
Tipp: Schauen Sie sich Rechenregeln für das Rechnen mit Varianzen an.

(c) Testen Sie, ob die Formel

$$S_0^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 \tag{3}$$

eine erwartungstreue Schätzfunktion für die Varianz σ^2 der Grundgesamtheit ist. Falls die Schätzfunktion nicht erwartungstreu ist, suchen Sie nach einer geeigneten Korrektur.

(d) Meist ist die Varianz σ^2 der Grundgesamtheit unbekannt und es wird der Schätzer (1) für μ genutzt und (3) wird zu:

$$S_1'^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2. \tag{4}$$

Testen Sie, ob (4) eine erwartungstreue Schätzfunktion für die Varianz σ^2 der Grundgesamtheit ist. Falls die Schätzfunktion nicht erwartungstreu ist, suchen Sie nach einer geeigneten Korrektur.
Tipp: Erweitern Sie den Summanden mit $-\mu+\mu$ und nutzen Sie die gegebene Relation (2).

Nico Guth
Jan Philipp Jäkel
David Venker

a) $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

$E(\hat{\mu}) = E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$

$= \frac{1}{n} E\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i)$

$= \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{1}{n}\cdot n\cdot\mu = \mu$ erwartungstreu

b) zu zeigen: $E((\bar{X}-\mu)^2) = Var(\bar{X}) = \frac{\sigma^2}{n}$

$Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right)$

$= \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{\sigma^2}{n}$

mit $Var(X_i) = \sigma^2$

c) $S_0^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$

$E(\hat{\sigma}^2) = E(S_0^2) = E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\right)$

$= \frac{1}{n}\sum_{i=1}^{n} E\left((X_i-\mu)^2\right) = \frac{n}{n}\sigma^2 = \sigma^2$

d) $S_1'^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$

$E(\hat{\sigma}^2) = E(S_1'^2) = \frac{1}{n}\sum_{i=1}^{n} E\left((X_i-\bar{X}+\mu-\mu)^2\right) = \frac{1}{n}\sum_{i=1}^{n} E\left(\left[(X_i-\mu)+(\mu-\bar{X})\right]^2\right)$

$$E(\hat{\sigma}^2) = E(S_1'^2) = \frac{1}{n}\sum_{i=1} E\left((X_i - \bar{X} + \mu - \mu)^2 \right) = \frac{1}{n}\sum_{i=1}^{n} E\left(\left[(X_i - \mu) + (\mu - \bar{X}) \right]^2 \right)$$

$$= \frac{1}{n}\sum_{i=1}^{n} E\left((X_i - \mu)^2 + 2(X_i - \mu)(\mu - \bar{X}) + (\mu - \bar{X})^2 \right)$$

$$= \sigma^2 + \frac{1}{n}\sum_{i=1}^{n} E\left(2(X_i - \mu)(\mu - \bar{X}) + (\mu - \bar{X})^2 \right) \qquad \left[\frac{1}{n}\sum_{i=1}^{n} E\left((X_i - \mu)^2\right) = \sigma^2 \right]$$

$$= \sigma^2 + \frac{1}{n} E\left(-\sum_{i=1}^{n} 2(X_i - \mu)(\bar{X} - \mu) + \sum_{i=1}^{n} (\bar{X} - \mu)^2 \right) \qquad \left[E\left((\bar{X} - \mu)^2\right) = \frac{\sigma^2}{n} \right]$$

$$= \sigma^2 + \frac{\sigma^2}{n} - \frac{1}{n} E\left(\sum_{i=1}^{n} 2(X_i - \mu)(\bar{X} - \mu) \right), \qquad \left[\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \right]$$

$$= \sigma^2 + \frac{\sigma^2}{n} - \frac{2}{\cancel{n}} E\left((\bar{X} - \mu) \cdot \cancel{n}(\bar{X} - \mu) \right) \qquad \left[\sum_{i=1}^{n} (X_i - \mu) = \left(\sum_{i=1}^{n} X_i \right) - n\mu = n \cdot (\bar{X} - \mu) \right]$$

$$= \sigma^2 + \frac{\sigma^2}{n} - 2 \cdot E\left((\bar{X} - \mu)^2 \right)$$

$$= \frac{n\sigma^2}{n} + \frac{\sigma^2}{n} - 2 \cdot \frac{\sigma^2}{n}$$

$$= \frac{n\sigma^2}{n} - \frac{\sigma^2}{n} = \frac{\sigma^2 (n-1)}{n}$$

Der Faktor $\frac{(n-1)}{n}$ ist eine Verzerrung!

Also: $\frac{n}{(n-1)} \hat{S}_1^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2 = S_1^2$

$$\frac{n}{n-1} E(S_1'^2) = \sigma^2 = E\left(\frac{n}{n-1} S_1'^2\right) \quad \checkmark$$

blatt13_guth_venker_jaekel

October 31, 2021

1 Aufgabe 30 Methode der kleinsten Quadrate

[1]:
```
import numpy as np
import matplotlib.pyplot as plt
```

1.1 a) Designmatrix A

hier Asymmetrie $= f(\Psi) = a_1 f_1(\Psi) + a_2 f_2(\Psi)$
mit $f_1(\Psi) = \cos(\Psi)$, $f_2(\Psi) = \sin(\Psi)$
Designmatrix: $A = \begin{pmatrix} f_1(\Psi_1) & f_2(\Psi_1) & \dots \\ f_1(\Psi_2) & f_2(\Psi_2) & \dots \\ \dots & \dots & \dots \end{pmatrix}$

[2]:
```
psi = np.arange(0,331,30) # in Grad
psi = np.radians(psi) # in rad
asym = [-0.032,0.010,0.057,
        0.068,0.076,0.080,
        0.031,0.005,-0.041,
        -0.090,-0.088,-0.074]
asym = np.array(asym)
asym_err = 0.011

print(f'{psi.shape = }')
print(f'{asym.shape = }')
```

```
psi.shape = (12,)
asym.shape = (12,)
```

[3]:
```
# design matrix:
A = np.column_stack([np.cos(psi),np.sin(psi)])

print('Designmatrix A:')
print(f'{A.shape = }')
print(f'{A = }')
```

```
Designmatrix A:
A.shape = (12, 2)
A = array([[ 1.00000000e+00,  0.00000000e+00],
       [ 8.66025404e-01,  5.00000000e-01],
```

```
       [ 5.00000000e-01,  8.66025404e-01],
       [ 6.12323400e-17,  1.00000000e+00],
       [-5.00000000e-01,  8.66025404e-01],
       [-8.66025404e-01,  5.00000000e-01],
       [-1.00000000e+00,  1.22464680e-16],
       [-8.66025404e-01, -5.00000000e-01],
       [-5.00000000e-01, -8.66025404e-01],
       [-1.83697020e-16, -1.00000000e+00],
       [ 5.00000000e-01, -8.66025404e-01],
       [ 8.66025404e-01, -5.00000000e-01]])
```

1.2 b) Lösungsvektor $\hat{\vec{a}}$

$\hat{\vec{a}} = (A^T A)^{-1} A^T \vec{y}$
mit $\vec{y} =$ Asymmetrien

[4]:
```python
y = asym
a = np.linalg.inv(A.T @ A) @ A.T @ y

print(f'{a.shape = }')
print(f'{a = }')
```

```
a.shape = (2,)
a = array([-0.0375063 ,  0.07739978])
```

1.3 c) Kovarianzmatrix $V[\hat{\vec{a}}]$

$V[\vec{y}] = \sigma^2 1$
$V[\hat{\vec{a}}] = \sigma^2 (A^T A)^{-1}$

[5]:
```python
sigma = asym_err
```

[6]:
```python
# Kovarianzmatrix:
V_a = sigma**2 * np.linalg.inv(A.T @ A)

print('Kovarianzmatrix von a:')
print(f'{V_a.shape = }')
print(f'{V_a = }')
```

```
Kovarianzmatrix von a:
V_a.shape = (2, 2)
V_a = array([[ 2.01666667e-05, -1.67921232e-21],
       [-1.67921232e-21,  2.01666667e-05]])
```

[7]:
```python
# Fehler von a:
a_err = np.sqrt(np.diag(V_a))
# Korrelationskoeffizient:
rho = V_a[1,0] / a_err[0]**2
```

```
print(f'Fehler von a = {a_err}')
print(f'Korrelationskoeffizient von a = {rho}')
```

```
Fehler von a = [0.00449073 0.00449073]
Korrelationskoeffizient von a = -8.326672684688675e-17
```

1.4 d) A_0 und δ berechnen (aus a_1 und a_2, mit Fehler und Korrelation)

$f(\Psi) = A_0 \cos(\Psi + \delta) = a_1 \cos(\Psi) + a_2 \sin(\Psi)$
siehe OneNote weiter unten für folgendes:
$A_0 = \frac{a_1}{\cos(\arctan(-a_2/a_1))}$
$\delta = \arctan(-a_2/a_1)$

$$V\left[\begin{pmatrix} A_0 \\ \delta \end{pmatrix}\right] = J \cdot V[\hat{\vec{a}}] \cdot J^T$$

mit $J = \begin{pmatrix} \left(\frac{a_2^2}{a_1^2}+1\right)^{-1/2} & \frac{a_2}{a_1}\left(\frac{a_2^2}{a_1^2}+1\right)^{-1/2} \\ \frac{a_2}{a_1^2+a_2^2} & -\frac{a_1}{a_1^2+a_2^2} \end{pmatrix}$

[8]:
```
# A_0 und delta:
a1 = a[0]
a2 = a[1]
A_0 = a1 / np.cos(np.arctan(-a2/a1))
delta = np.arctan(-a2/a1)

print(f'{A_0 = }')
print(f'{delta = }')
```

```
A_0 = -0.0860084162956094
delta = 1.1195615399310932
```

[9]:
```
# Jacobi-Matrix:
J = [[(a2**2/a1**2+1)**-(1/2), (a2/a1)*(a2**2/a1**2+1)**-(1/2)],
     [a2/(a1**2+a2**2), -a1/(a1**2+a2**2)]]
J = np.array(J)

print('Jacobi-Matrix:')
print(f'{J = }')
```

```
Jacobi-Matrix:
J = array([[ 0.43607706, -0.89990933],
       [10.46303798,  5.07016733]])
```

[10]:
```
# Kovarianzmatrix:
V = J @ V_a @ J.T

print('Kovarianzmatrix:')
print(f'{V = }')
```

```
Kovarianzmatrix:
V = array([[2.01666667e-05, 2.71050543e-20],
       [1.35525272e-20, 2.72616550e-03]])
```

[11]:
```
# Fehler von a:
err = np.sqrt(np.diag(V))
# Korrelationskoeffizient:
rho_ = V[1,0] / err[0]**2

print(f'Fehler von A_0 = {err[0]}')
print(f'Fehler von delta = {err[1]}')
print(f'Korrelationskoeffizient = {rho_}')
```

```
Fehler von A_0 = 0.004490731195102494
Fehler von delta = 0.05221269485613977
Korrelationskoeffizient = 6.720261399703538e-16
```

1.5 Endergebnis:

für $f(\Psi) = A_0 \cos(\Psi + \delta) = a_1 \cos(\Psi) + a_2 \sin(\Psi)$
ergibt sich aus den gegebenen Daten mit der Methode der kleinsten Quadrate:
$a_1 = (-0.0375 \pm 0.0045)$,
$a_2 = (0.0774 \pm 0.0045)$,
$A_0 = (-0.0860 \pm 0.0045)$,
$\delta = (1.120 \pm 0.052)$

Die Korrelationskoeffizienten sind mit $\approx 10^{-20}$ nahezu 0 und würden wir nicht weiter für eine Analyse beachten.
Zudem handelt es sich hierbei vermutlich eher um Rundungsfehler, da für A_0 und δ die Kovarianzmatrix nicht symmetrisch ist.

[ ]:
```
```

(d) Berechnen Sie A_0 und δ und deren Fehler und Korrelation aus a_1 und a_2.

$$f(\Psi) = A_0 \cos(\Psi + \delta) \stackrel{!}{=} a_1 \cos(\Psi) + a_2 \sin(\Psi)$$

mit $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$

$$\Rightarrow A_0 \cos(\Psi + \delta) = A_0 \cos(\Psi)\cos(\delta) - A_0 \sin(\Psi)\sin(\delta)$$

$$\stackrel{!}{=} a_1 \cos(\Psi) + a_2 \sin(\Psi) \quad \text{Koeffizientenvergleich}$$

$$\Rightarrow A_0 \cos\delta = a_1 \quad \text{und} \quad -A_0 \sin\delta = a_2$$

$$\Leftrightarrow A_0 = \frac{a_1}{\cos\delta} \quad \text{einsetzen}$$

$$\Rightarrow -\frac{a_1}{\cos\delta} \sin\delta = a_2$$

$$\Leftrightarrow \delta = \arctan\left(-\frac{a_2}{a_1}\right)$$

$$\Rightarrow A_0 = \frac{a_1}{\cos\left(\arctan\left(-\frac{a_2}{a_1}\right)\right)}$$

- Fehler und Korrelation:

$$\vec{y} = \vec{g}(\vec{x}) \Rightarrow Var[\vec{y}] = J \cdot Var[\vec{x}] \cdot J^T$$

mit $J \hat{=}$ Jacobi-Matrix von $\vec{g}(\vec{x})$

hier $\vec{y} = \begin{pmatrix} A_0 \\ \delta \end{pmatrix}$ $\quad \vec{x} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ $\quad \vec{g}(\vec{x}) = \begin{pmatrix} \frac{a_1}{\cos\left(\arctan\left(-\frac{a_2}{a_1}\right)\right)} \\ \arctan\left(-\frac{a_2}{a_1}\right) \end{pmatrix}$

$$J = \begin{pmatrix} \frac{\partial g_1}{\partial a_1} & \frac{\partial g_1}{\partial a_2} \\ \frac{\partial g_2}{\partial a_1} & \frac{\partial g_2}{\partial a_2} \end{pmatrix}$$

$$\frac{\partial g_1}{\partial a_1} = \frac{\partial}{\partial a_1}\left(\frac{a_1}{\cos(\arctan(-\frac{a_2}{a_1}))}\right) = \left(\left(\frac{a_2}{a_1}\right)^2 + 1\right)^{-\frac{1}{2}}$$

$$\frac{\partial g_2}{\partial a_1} = \frac{a_2}{a_1^2 + a_2^2}$$

$$\frac{\partial g_1}{\partial a_2} = \frac{a_2}{a_1}\left(\left(\frac{a_2}{a_1}\right)^2 + 1\right)^{-\frac{1}{2}}$$

$$\frac{\partial g_2}{\partial a_2} = -\frac{a_1}{a_1^2 + a_2^2}$$

$$J = \begin{pmatrix} \left(\left(\frac{a_2}{a_1}\right)^2 + 1\right)^{-\frac{1}{2}} & \frac{a_2}{a_1}\left(\left(\frac{a_2}{a_1}\right)^2 + 1\right)^{-\frac{1}{2}} \\ \frac{a_2}{a_1^2 + a_2^2} & -\frac{a_1}{a_1^2 + a_2^2} \end{pmatrix}$$

$$Var\left[\begin{pmatrix} A_0 \\ \delta \end{pmatrix}\right] = J \cdot Var\left[\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}\right] \cdot J^T$$