SMD-Abgabe

11. Übungsblatt

Lars Kolk
lars.kolk@tu-dortmund.de

Julia Sobolewski
julia.sobolewski@tu-dortmund.de

Jannine Salewski
jannine.salewski@tu-dortmund.de

Abgabe: 24.01.2018

32	33	34	35	36	Σ
3/5	4/5	k.A 0/5	3,5/5	4,75/5	15,25/20

TU Dortmund – Fakultät Physik

Aufgabe 35: Likelihood-Quotienten-Test

a) $H_0: \mu = \mu_0 \quad \Rightarrow \Theta_0 = \{(\mu, \sigma^2): \mu = \mu_0, \sigma^2 > 0\}$ ✓

$H_1: \mu \neq \mu_0 \quad \Rightarrow \Theta_1 = \{(\mu, \sigma^2): \mu \neq \mu_0, \sigma^2 > 0\}$ ✓

$$\Rightarrow \mathcal{L}((\mu, \sigma^2)|x) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)$$

$$= \frac{1}{(\sqrt{2\pi\sigma^2})^n} \exp\left(-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right) \checkmark$$

Teststatistik: $\Gamma(x) = \frac{\sup_{\theta\in\Theta_0} \mathcal{L}(\theta|x)}{\sup_{\theta\in\Theta_1} \mathcal{L}(\theta|x)}$ ✓ 1/1 P.

b) $$\frac{\partial}{\partial\sigma} \log(\mathcal{L}) = \frac{\partial}{\partial\sigma}\left[\log\left(\frac{1}{(2\pi)^{n/2}} \cdot \sigma^{-n}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right] \overset{!}{=} 0$$

$$0 = (2\pi)^{n/2} \cdot \cancel{\sigma^{n}} \cdot (-n) \cdot (2\pi)^{-n/2} \cdot \sigma^{-n-1} + \frac{1}{\sigma^3} \cdot \sum_{i=1}^{n}(x_i-\mu)^2$$

$$0 = \frac{-n}{\sigma} + \frac{1}{\sigma^3} \sum_{i=1}^{n}(x_i-\mu)^2$$

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 \checkmark$$

$$\frac{\partial}{\partial\mu} \log(\mathcal{L}) = \frac{\partial}{\partial\mu}\left[\log\left(\frac{1}{(2\pi)^{n/2}} \sigma^{-n}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right] \overset{!}{=} 0$$

$$0 = \frac{-1}{2\sigma^2}\sum_{i=1}^{n} 2(x_i-\mu)\cdot(-1) = \frac{1}{\sigma^2}\sum_{i=1}^{n} x_i - n\mu$$

$$\Rightarrow \mu = \frac{1}{n}\sum_{i=1}^{n} x_i \checkmark$$

$H_0: \mu = \mu_0$ ✓

$H_1: \mu = \frac{1}{n}\sum_{i=1}^{n} x_i$ ✓ 1/1 P.

c) $$\sup_{\theta\in\Theta_0} \mathcal{L}(\theta|x) = \frac{1}{(2\pi\hat{\sigma}^2)^{n/2}} \exp\left(-\frac{1}{2\hat{\sigma}^2}\sum_{i=1}^{n}(x_i-\mu_0)^2\right)$$

$$= \frac{1}{\left(\frac{2\pi}{n}\sum_{i=1}^{n}(x_i-\mu_0)^2\right)^{n/2}} \underbrace{\exp\left(\frac{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu_0)^2}{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu_0)^2}\right)}_{\exp(-\frac{n}{2})}$$

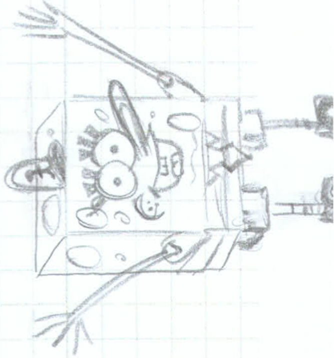
$$\sup_{\theta\in\Theta_1} \mathcal{L}(\theta|x) = \left(\frac{2\pi}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)^{-n/2} \exp\left(-\frac{n}{2}\right)$$

$$\Rightarrow \Gamma(x) = \left(\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\mu_0)^2}\right)^{n/2} \checkmark$$

für eine Stichprobenvarianz gilt $s^2 = \frac{1}{n\cancel{-1}} \sum_{i=1}^{n} (x_i - \bar{x})^2$

$$\Rightarrow \Gamma(x) = \left(\frac{s^2(n-1)}{\sum_{i=1}^{n} (x_i - \mu_0)^2} \right)^{n/2} = \left(\frac{s^2(n-1)}{n(\bar{x} - \mu_0)^2} \right)^{n/2} < C$$

1/2 P.



d) $\sqrt{N} = \sqrt{25} = 5$ ✓ $\mu_0 = 200ml$ ✓ $\bar{x} = 205ml$ ✓ $s = 10ml$ ✓

$$\Rightarrow T = 5 \cdot \frac{(205ml - 200ml)}{10ml} = 2{,}5$$ ✓

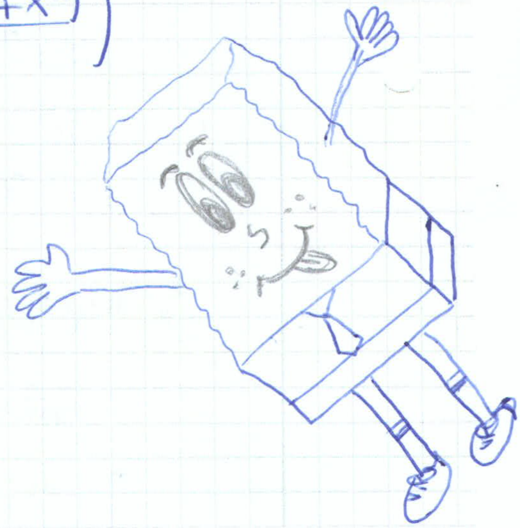
abgelehnt oder blibehalten ?

0,5/1 P.

$$\Gamma(x) = \exp\left(- \frac{\sum (x_i^2 - 2\mu_0 x_i + \mu_0^2) - \sum (x_i^2 - 2\bar{x}x_i + \bar{x}^2)}{2s^2} \right)$$

$$= \exp\left(- \frac{n2\bar{x}(\bar{x} - \mu_0) + n\mu_0^2 - n\bar{x}^2}{2s^2} \right)$$

$$= \exp\left(- \frac{n(\bar{x} - \mu_0)^2}{2s^2} \right)$$



$\Rightarrow \Gamma(x)$ monoton fallend in $t = \frac{\sqrt{n}(x - \mu_0)}{s}$

e) $t_{24; 1 - \frac{0{,}05}{2}} = 2{,}064$

Aufgabe 32

In [74]:
```
import numpy as np
from scipy.stats import chi2
```

In [75]:
```
def chi_quadrat_test(data, sigma, hypothese, freiheitsgrade, signifi
kanz):
    pruefgroesse = sum((data-hypothese)**2/sigma**2)
    chi_quadrat = chi2.ppf(1-signifikanz, freiheitsgrade)
    test = pruefgroesse < chi_quadrat

    print('Prüfgröße: {:.2f} \nChi_quadrat: {:.2f} \nTest: {}'.forma
t(pruefgroesse, chi_quadrat, test))
```

müsste in der Klausur per Hand gehen :)

a)

In [76]:
```
data = np.array([31.6, 32.2, 31.2, 31.9, 31.3, 30.8, 31.3])
sigma = 0.5
fg = len(data)-1

H_A = 31.3
signifikanz = 0.05
```

f

es muss kein Freiheitsgrad abgezogen werden, da kein Parameter geschätzt wird => dof = 7

In [77]:
```
chi_quadrat_test(data, sigma, H_A, fg, signifikanz)
```

```
Prüfgröße: 6.08
Chi_quadrat: 12.59
Test: True
```

✓ f 14,07 (✓)

1,5/3P.

b)

In [78]:
```
H_B = 30.7

chi_quadrat_test(data, sigma, H_B, fg, signifikanz)
```

```
Prüfgröße: 21.92
Chi_quadrat: 12.59
Test: False
```

✓ f 14,07 (✓)

1,5/2

Was bedeutet True/False? Abgelehnt oder nicht abgelehnt? nicht eindeutig, aber kein Punktabzug

Aufgabe 32

$\epsilon_A = 31{,}3\,meV$; $\epsilon_B = 30{,}7\,meV$; Freiheitsgrade = 7 ; $\alpha = 0{,}05$

$$\chi^2 = \sum_{i=1}^{n} \frac{(\epsilon_i - \epsilon_A)^2}{\sigma_i^2}$$

$$\chi^2(Fg = 7) = 14{,}07$$

Aufgabe 33

a)

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

Poisson: $P(x) \approx \frac{1}{\sqrt{2\pi\lambda}} e^{\frac{-(x-\lambda)^2}{2\lambda}}$ mit $\lambda \geq 30, \lambda = \mu = \sigma^2$

0,5/0,5

b)

In [1]:
```
import numpy as np
import matplotlib.pyplot as plt
from scipy import stats
np.random.seed(1337)
```

In [2]:
```
def ktest(X, Y, alpha):
    n=np.sum(X)
    m=np.sum(Y)

    FX=np.cumsum(X)/n
    FY=np.cumsum(Y)/m

    d = np.max(np.abs(FX-FY)) #maximale Abstand
    #print(d)
    K_alpha = np.sqrt(np.log(2/alpha)/2)
    #print(K_alpha)
    passed = True

    #Überprüfung der Hypothese
    if np.sqrt((n*m)/(n+m))*d <= K_alpha :
        return True
    else:
        return False
```

c) + d)

kleinste Quadrate!
→ KLAUSUR!
8 Aufgaben
3h
Din A5 Zettel

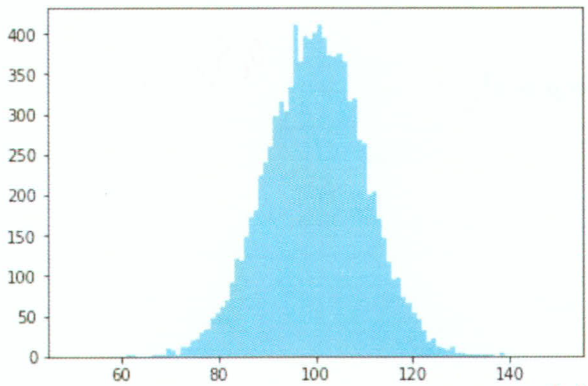
muss in die Schleife, mu & sigma ändern sich

In [3]:
```
my=100
sig=10
normal = np.random.normal(100, 10, 10000)
normal.astype(int)

normal = np.delete(normal, np.where(normal < my-5*sig))
normal = np.delete(normal, np.where(normal > my+5*sig))


normal_bins = np.histogram(normal, bins=np.linspace(my-5*sig,my+5*sig,100))


plt.hist(normal, bins=np.linspace(my-5*sig,my+5*sig,100), alpha=0.5)
plt.show()
```



Warum erst ab 30???

In [4]:
```
for alpha in [0.05, 0.025, 0.01]:
    for lam in np.linspace(30, 100, 10000):
        pois=np.random.poisson(lam, 10000)

        pois=np.delete(pois, np.where(pois < my-5*sig))
        pois=np.delete(pois, np.where(pois > my+5*sig))

        poisson_bins = np.histogram(pois, bins=np.linspace(my-5*sig,my+5*sig,100
))

        if ktest(poisson_bins[0], normal_bins[0], alpha) == True:
            print("alpha = ", alpha, " : Für den Test sind die Verteilungen ab L
ambda = ", lam, "ununterscheidbar. \n" )
            break
```

```
alpha =  0.05  : Für den Test sind die Verteilungen ab Lambda =  99.6219621962
1963 ununterscheidbar.

alpha =  0.025  : Für den Test sind die Verteilungen ab Lambda =  30.112011201
12011 ununterscheidbar.

alpha =  0.01  : Für den Test sind die Verteilungen ab Lambda =  30.0420042004
20042 ununterscheidbar.
```

FF

Ø solltet ihr nehmen

In [ ]:

b) 1/2P

c)/d) 2,5/2,5

Aufgabe 36

$$\sum_{i=1}^{r} p_i = 1, \qquad N = \sum_{i=1}^{r} n_i, \qquad M = \sum_{i=1}^{r} m_i$$

a)

Die Zählraten der einzelnen Bins folgen jeweils einer Poisson-Verteilung:

$$f(n_i) = \frac{(Np_i)^{n_i}}{n_i!} e^{-Np_i}$$

$$g(m_i) = \frac{(Mp_i)^{m_i}}{m_i!} e^{-Mp_i}$$

1/1P.

b)

Aufstellen der Likelihood-Funktion:

$$\mathcal{L} = \frac{(Np_i)^{n_i}}{n_i!} e^{-Np_i} \cdot \frac{(Mp_i)^{m_i}}{m_i!} e^{-Mp_i} = \frac{N^{n_i} M^{m_i}}{n_i! m_i!} \cdot p_i^{n_i+m_i} \cdot e^{-(N+M)p_i}$$

Berechnung des Maximums $\hat{p}_i$:

$$\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\hat{p}_i} = \frac{N^{n_i} M^{m_i}}{n_i! m_i!} \left[(n_i + m_i) \hat{p}_i^{n_i+m_i-1} e^{-(N+M)\hat{p}_i} - \hat{p}_i^{n_i+m_i} (N+M) e^{-(N+M)\hat{p}_i} \right]$$

$$= \frac{N^{n_i} M^{m_i}}{n_i! m_i!} \hat{p}_i^{n_i+m_i} e^{-(N+M)\hat{p}_i} \left(\frac{n_i + m_i}{\hat{p}_i} - (N+M) \right) \stackrel{!}{=} 0$$

$$\Rightarrow \hat{p}_i = \frac{n_i + m_i}{M + N}$$

Maximum ?
=> zweite Ableitung ?
noch 1/1P.

c)

$$\underbrace{\chi^2 = \sum_{i=1}^{r} \frac{(n_i - N\hat{p}_i)^2}{N\hat{p}_i} + \sum_{i=1}^{r} \frac{(m_i - M\hat{p}_i)^2}{M\hat{p}_i}}_{\text{Gaußnäherung} \to \text{gilt nur für ausreichend große Statistik}}$$

1/1P.

d)

Die Anzahl der Freiheitsgrade beträgt $r - 1$.

Die Teststatistik folgt für kleine Bininhalte keiner χ^2-Verteilung mehr, weil...

0,75/1

e)

In [14]:

```
import numpy as np
from scipy.stats import chi2
```

In [15]:
```
n = np.array([111, 188, 333])
m = np.array([15, 36, 30])

r = len(n)-1
N = sum(n)
M = sum(m)

def chi_quadrat_test(n, m, freiheitsgrade, signifikanz):
    pruefgroesse = 1/(M*N) * sum((N*m-M*n)**2/(n+m))
    chi_quadrat = chi2.ppf(1-signifikanz, freiheitsgrade)
    test = pruefgroesse < chi_quadrat

    print('Prüfgröße: {:.2f} \nChi_quadrat: {:.2f} \nTest: {}'.format(pruefgroes
se, chi_quadrat, test))
```

In [16]:
```
chi_quadrat_test(n, m, r, 0.1)
```

```
Prüfgröße: 8.43 ✓
Chi_quadrat: 4.61 ✓
Test: False
```

In [17]:
```
chi_quadrat_test(n, m, r, 0.05)
```

```
Prüfgröße: 8.43 ✓
Chi_quadrat: 5.99 ✓
Test: False
```

In [18]:
```
chi_quadrat_test(n, m, r, 0.01)
```

```
Prüfgröße: 8.43 ✓
Chi_quadrat: 9.21 ✓
Test: True
```

Die Nullhypothese wird nur für $\alpha = 0,01$ nicht verworfen. ✓

1/1P.