

Topics covered in this tutorial

- Conditional densities, blind and MMSE estimation
 - ML, MAP and MMSE estimation in Gaussian noise
 - Variance estimators, bias, consistency and precision
 - Hypothesis tests, error probabilities and decision costs
 - Bonus: Chernoff bounds and alarm-based MMSE
- Complete all main questions (1–5) before the tutorial test that follows. Bonus Questions 6 and 7 are optional. The test normally starts between approximately 16:00 and 16:15.*

Onderwerpe wat in hierdie tutoriaal gedek word

- Voorwaardelike digthede, blinde en MMKF-skatting
- MA-, MAP- en MMKF-skatting in Gaussiese ruis
- Variansieberamers, sydigheid, konsekwentheid en presisie
- Hipotesetoetse, foutwaarskynlikhede en besluitkoste
- Bonus: Chernoff-grense en alarmgebaseerde MMKF

Voltooi al die hoofvrae (1–5) voor die tutoriaaltoets wat daarop volg. Bonusvrae 6 en 7 is opsioneel. Die toets begin gewoonlik tussen ongeveer 16:00 en 16:15.

1. The random variables X and Y have joint PDF

$$f_{X,Y}(x,y) = \begin{cases} 3(1+x-y), & 0 \leq x \leq y \leq 1, \\ 0, & \text{otherwise / andersins.} \end{cases}$$

- Find $f_{X|Y}(x|y)$ for $0 < y \leq 1$.
- Find the MMSE estimate $\hat{x}_M(y)$ of X given $Y = y$.
- Find $f_{Y|X}(y|x)$ for $0 \leq x < 1$.
- Find the MMSE estimate $\hat{y}_M(x)$ of Y given $X = x$.
- Find the blind MMSE estimate of X . Explain why observing Y can improve estimation compared with always using this constant.

2. A signal X is transmitted over a noisy channel. The received signal is $Y = X + W$, where X and W are independent Gaussian random variables with means zero and positive variances σ_X^2 and σ_W^2 , respectively. The value $Y = y$ is observed.

- Find the maximum-likelihood (ML) estimate of X .
- Find the maximum a posteriori (MAP) estimate of X .
- Calculate $E[Y]$, $\text{Var}[Y]$ and $\text{Cov}[X, Y]$. Hence find the MMSE estimate $\hat{x}_M(y)$ of X using the Gaussian conditional-mean formula.

Hint: X and Y are jointly Gaussian; the MMSE estimate is $E[X | Y = y]$.

1. Die toevalsveranderlikes X en Y het gesamentlike WDF

- Vind $f_{X|Y}(x|y)$ vir $0 < y \leq 1$.
- Vind die MMKF-skatting $\hat{x}_M(y)$ van X gegewe $Y = y$.
- Vind $f_{Y|X}(y|x)$ vir $0 \leq x < 1$.
- Vind die MMKF-skatting $\hat{y}_M(x)$ van Y gegewe $X = x$.
- Vind die blinde MMKF-skatting van X . Verduidelik waarom die waarneming van Y die skatting kan verbeter teenoor die gebruik van hierdie konstante.

2. 'n Sein X word oor 'n ruiserige kanaal versend. Die ontvangde sein is $Y = X + W$, waar X en W onafhanklike Gaussiese toevalsveranderlikes is met gemiddeldes nul en positiewe variansies σ_X^2 en σ_W^2 , onderskeidelik. Die waarde $Y = y$ word waargeneem.

- Vind die maksimum-aanneemlikheidskatting (MA) van X .
- Vind die maksimum a posteriori-skatting (MAP) van X .
- Bereken $E[Y]$, $\text{Var}[Y]$ en $\text{Cov}[X, Y]$. Vind daaruit die MMKF-skatting $\hat{x}_M(y)$ van X met die Gaussiese formule vir die voorwaardelike gemiddelde. *Wenk: X en Y is gesamentlik Gaussies; die MMKF-skatting is $E[X | Y = y]$.*

3. A digital scale repeatedly measures an object whose true weight is 50 kg. Let X_i be the i th measured weight minus 50 kg. The errors are iid Gaussian random variables with known mean zero and unknown variance $\sigma^2 < 0.5 \text{ kg}^2$. Use the known-mean variance estimator below; its variance is also given.

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2,$$

- (a) Show that $\hat{\sigma}_n^2$ is an unbiased estimator of σ^2 .
- (b) Find the smallest number of measurements that guarantees a standard error below 0.08 kg^2 .
- (c) Use Chebyshev's inequality to show that $\hat{\sigma}_n^2$ is consistent: for each fixed $\varepsilon > 0$, the probability that its error is at least ε tends to zero as $n \rightarrow \infty$.

4. A transmitter sends a voltage X with support $S_X = \{0, 4\}$ volts. Define $H_0 = \{X = 0\}$ and $H_1 = \{X = 4\}$, with $P(H_0) = 0.8$ and $P(H_1) = 0.2$. A receiver reports only a low (L), medium (M) or high (H) reading, so $S_Y = \{L, M, H\}$. The conditional probabilities are given below.

y	$P(Y = y H_0)$	$P(Y = y H_1)$
L	0.6	0.1
M	0.3	0.4
H	0.1	0.5

- (a) Calculate $P_Y(y)$ and $P(H_1 | Y = y)$ for each of the three readings. Present your results in a table.
- (b) State the ML and MAP decisions for each reading. Identify any reading for which the two rules disagree.
- (c) For the MAP rule, calculate the false-alarm probability (choose H_1 when H_0 is true), the miss probability (choose H_0 when H_1 is true), and the overall probability of error.
- (d) Now a false alarm costs 1 unit and a miss costs 4 units; correct decisions cost zero. Find the minimum-cost decision for each reading. Compare its expected cost with the expected cost of the MAP rule.
- (e) Find the blind MMSE estimate of X and the MMSE estimate for each reading. Explain why an MMSE estimate need not be 0 or 4 volts.

3. 'n Digitale skaal meet herhaaldelik 'n voorwerp met ware gewig 50 kg. Laat X_i die i de gemete gewig minus 50 kg wees. Die foute is oiv Gaussiese toevalsveranderlikes met bekende gemiddelde nul en onbekende variansie $\sigma^2 < 0.5 \text{ kg}^2$. Gebruik die variansieberamer met bekende gemiddelde hieronder; die variansie daarvan word ook gegee.

$$\text{Var}[\hat{\sigma}_n^2] = \frac{2\sigma^4}{n}.$$

- (a) Toon dat $\hat{\sigma}_n^2$ 'n onsydige beramer van σ^2 is.
- (b) Vind die kleinste aantal metings wat 'n standaardfout onder 0.08 kg^2 waarborg.
- (c) Gebruik Chebyshev se ongelykheid om te toon dat $\hat{\sigma}_n^2$ konsekwent is: vir elke vaste $\varepsilon > 0$ neig die waarskynlikheid dat die fout minstens ε is na nul soos $n \rightarrow \infty$.

4. 'n Sender stuur 'n spanning X met steun $S_X = \{0, 4\}$ volt. Definieer $H_0 = \{X = 0\}$ en $H_1 = \{X = 4\}$, met $P(H_0) = 0.8$ en $P(H_1) = 0.2$. 'n Ontvanger rapporteer slegs 'n lae (L), medium (M) of hoë (H) lesing, dus $S_Y = \{L, M, H\}$. Die voorwaardelike waarskynlikhede word hieronder gegee.

- (a) Bereken $P_Y(y)$ en $P(H_1 | Y = y)$ vir elk van die drie lesings. Gee jou resultate in 'n tabel.
- (b) Gee die MA- en MAP-besluite vir elke lesing. Identifiseer enige lesing waarvoor die twee reëls verskil.
- (c) Bereken vir die MAP-reël die valsalarmwaarskynlikheid (kies H_1 wanneer H_0 waar is), die miswaarskynlikheid (kies H_0 wanneer H_1 waar is), en die algehele foutwaarskynlikheid.
- (d) Nou kos 'n valsalarm 1 eenheid en 'n mis 4 eenhede; korrekte besluite kos nul. Vind die minimumkostebesluit vir elke lesing. Vergelyk die verwagte koste daarvan met die verwagte koste van die MAP-reël.
- (e) Vind die blinde MMKF-skatting van X en die MMKF-skatting vir elke lesing. Verduidelik waarom 'n MMKF-skatting nie noodwendig 0 of 4 volt hoef te wees nie.

5. Independent temperature readings $T_1, \dots, T_n$ come from the same distribution, with unknown mean μ and finite variance σ^2 . Define the sample mean and two variance estimators as follows.

$$M_n = \frac{1}{n} \sum_{i=1}^n T_i, \quad V_n = \frac{1}{n} \sum_{i=1}^n (T_i - M_n)^2, \quad V'_n = \frac{1}{n-1} \sum_{i=1}^n (T_i - M_n)^2, \quad n \geq 2.$$

(a) For the five readings 21, 22, 22, 23, 27 in degrees Celsius, calculate the observed values of M_5 , V_5 and V'_5 .

(b) Using the unbiased variance estimate V'_5 , estimate the standard error of the sample mean M_5 and the standard deviation of an individual reading. Briefly explain the difference between these two quantities.

Bonus questions

Optional extra practice.

6. The number X of requests received by a server in a fixed observation window is Poisson with mean 20. Its MGF is

$$\phi_X(s) = \exp\{20(e^s - 1)\}.$$

(a) Use the Chernoff bound to obtain an upper bound on $P(X \geq 30)$. Show how you choose the value of $s \geq 0$ that minimises the bound.

(b) Calculate the Markov bound for the same event and compare the two bounds. Explain why neither result is the actual probability of the event.

7. The temperature T in an enclosure is uniformly distributed over $[20, 30]$ degrees Celsius. A sensor reports only whether the alarm is on or off: it is on for $B = \{T > 26\}$ and off for $\bar{B}$. The observer does not see the exact temperature.

(a) Without an alarm reading, find the blind MMSE estimate of T and its mean square error.

(b) Calculate $P(B)$ and $P(\bar{B})$, and find the conditional PDFs $f_{T|B}(t)$ and $f_{T|\bar{B}}(t)$.

(c) Find the MMSE estimate when the alarm is on, and when it is off.

(d) Calculate the conditional mean square error for each alarm state. Hence find the average mean square error of this alarm-based estimator and compare it with the blind estimator.

5. Onafhanklike temperatuurlesings $T_1, \dots, T_n$ kom uit dieselfde verdeling, met onbekende gemiddelde μ en eindige variansie σ^2 . Definieer die steekproefgemiddelde en twee variansieberamers soos volg.

(a) Bereken vir die vyf lesings 21, 22, 22, 23, 27 in grade Celsius die waargenome waardes van M_5 , V_5 en V'_5 .

(b) Gebruik die onsydige variansieberaming V'_5 om die standaardfout van die steekproefgemiddelde M_5 en die standaardafwyking van 'n individuele lesing te beraam. Verduidelik kortliks die verskil tussen hierdie twee groothede.

Bonusvrae

Opsionele ekstra oefening.

6. Die aantal versoeke X wat 'n bediener in 'n vaste waarnemingsvenster ontvang, is Poisson met gemiddelde 20. Die MGF daarvan is

(a) Gebruik die Chernoff-grens om 'n bogrens vir $P(X \geq 30)$ te verkry. Toon hoe jy die waarde van $s \geq 0$ kies wat die grens minimaliseer.

(b) Bereken die Markov-grens vir dieselfde gebeurtenis en vergelyk die twee grense. Verduidelik waarom nie een van die resultate die werklike waarskynlikheid van die gebeurtenis is nie.

7. Die temperatuur T in 'n omhulsel is uniform verdeel oor $[20, 30]$ grade Celsius. 'n Sensor rapporteer slegs of die alarm aan of af is: dit is aan vir $B = \{T > 26\}$ en af vir $\bar{B}$. Die waarnemer sien nie die presiese temperatuur nie.

(a) Vind sonder 'n alarmlesing die blinde MMKF-skatting van T en die gemiddelde vierkantsfout daarvan.

(b) Bereken $P(B)$ en $P(\bar{B})$, en vind die voorwaardelike WDFs $f_{T|B}(t)$ en $f_{T|\bar{B}}(t)$.

(c) Vind die MMKF-skatting wanneer die alarm aan is, en wanneer dit af is.

(d) Bereken die voorwaardelike gemiddelde vierkantsfout vir elke alarmtoestand. Vind daaruit die gemiddelde vierkantsfout van hierdie alarmgebaseerde beramer en vergelyk dit met die blinde beramer.