# Probability Theory
# Exam July 2008 - Solutions

## Question 1

Let $A$ and $B$ be events in a probability space. $P(A) = 0.6$ and $P(B) = 0.2$. Furthermore, $B \subset A$. The probabilities $P(A^c \cup B)$ and $P(A^c \cup B^c)$ equal respectively...

## Solution

We can find $P(A^c \cup B)$ using

$$P(A^c \cup B) = P(A^c) + P(B) - P(A^c \cap B). \tag{1}$$

We have $P(A^c) = 1 - P(A) = 0.4$. Also, since $B \subset A$ we have $A^c \cap B = \emptyset$ (if you don't follow this, try drawing a Venn diagram), and thus $P(A^c \cap B) = P(\emptyset) = 0$. It follows that

$$P(A^c \cup B) = P(A^c) + P(B) = 0.4 + 0.2 = 0.6. \tag{2}$$

Now let's find $P(A^c \cup B^c)$. For that, we first note that (since $B \subset A$) we have

$$1 = P(A|B) = \frac{P(A \cap B)}{P(B)}, \tag{3}$$

so $P(A \cap B) = P(B) = 0.2$. It follows that

$$P(A^c \cup B^c) = 1 - P(A \cap B) = 1 - 0.2 = 0.8. \tag{4}$$

## Question 2

How many ways are there to choose a chairman, his deputy and a first and a second assistant from 30 participants at an election meeting?

## Solution

It is important to note that the chosen persons have different positions. So it DOES matter which person gets chosen first. We therefore need to find the amount permutations. This is

$$\frac{30!}{26!} = 30 \cdot 29 \cdot 28 \cdot 27 = 657720. \tag{5}$$

Another way to discover this is, is by using plain logics. We can choose our chairman out of 30 people. We can then choose his deputy out of 29 people, his first assistant out of 28 people and his second assistant out of 27 people. So the amount of ways in which we can do this is, indeed, $30 \cdot 29 \cdot 28 \cdot 27 = 657720$.

## Question 3

Three machines produce the same type of product in a factory. The first one gives 20% of the total production, the second one gives 30% and the third one 50%. It is known from past experience that 5%, 4% and 2% of the products made by machine 1, 2 and 3, respectively, are defective. If a single, randomly selected, product is defective, what is the probability that it was made by the first machine?

## Solution

We know that machine 1 makes 20% of the products, out of which 5% is defective. So $0.2 \cdot 0.05 = 0.01 = 1\%$ of the products are defective and made by machine 1. Similarly, 1.2% of the products are defective and made by machine 2. The percentage for machine 3 is 1%. So in total 3.2% of the products are defective. We know that 1% of the products are defective and made by the first machine. So that chance that a defective product is made by the first machine is

$$\frac{1}{3.2} = 0.3125. \tag{6}$$

Note that we have, in fact, applied Bayes' rule in the above computation.

## Question 4

Assume that the duration of horse pregnancies varies according to a normal distribution with mean 336 days and standard deviation 3 days. In what range of duration fall 99.7% of horse pregnancies (symmetrical with respect to the mean duration)?

## Solution

We can solve this using a TI calculator. We insert the equation $normalcdf(336 - X, 336 + X, 336, 3) = 0.997$ and let it solve for $X$. We find $X = 8.9$ days. So the range that was asked for is approximately $[327, 345]$.

## Question 5

Two aircraft fly at the same lateral position but at different altitudes. The altitude $\underline{h}_1$ of aircraft 1 is normally distributed with mean 8.0km and standard deviation 125m. The altitude $\underline{h}_2$ of aircraft 2 is normally distributed with mean 9.0km and standard deviation 125m. The random variables $\underline{h}_1$ and $\underline{h}_2$ are independent. What is the probability that the two aircraft are within a distance of 500m?

## Solution

Let's define the random variable $\underline{y} = \underline{h}_2 - \underline{h}_1$. It follows that $\overline{y} = 1000m$ and also that $\sigma_y^2 = \sigma_{h_1}^2 + \sigma_{h_2}^2$. It follows that $\overline{y} = \overline{h}_2 - \overline{h}_1 = 9000 - 8000 = 1000$ and also that $\sigma_y = 125\sqrt{2}$. We are now looking for the probability that $-500 \leq \underline{y} \leq 500$. We can find this using a TI calculator, and by inserting $normalcdf(-500, 500, 1000, 125\sqrt{2})$. We get as a result 0.00234.

## Question 6

We want to determine the mass of a block (i.e. a rectangular parallelepiped) made of copper by measuring the edge lengths $x_1$, $x_2$ and $x_3$ once. Before the actual measurement, we would like to assess the precision with which the block mass can be determined. We know that the edge length $x_1$ is approximately $10cm$, that the edge length $x_2$ is approximately $3cm$ and that the edge length $x_3$ is approximately $2cm$. We assume that the edge length measurements are independent and have a standard deviation (precision) of $1mm$. The density of copper is assumed to be infinitely precise and equals $8.9g/cm^3$. What is the standard deviation of the mass of the block, to a first-order approximation?

## Solution

We know that the mass is the density times the volume, so $\underline{M} = \rho \underline{V}$. (Note that we don't treat $\rho$ as a random variables because it is infinitely precise.) Also we have $\underline{V} = \underline{x}_1 \underline{x}_2 \underline{x}_3$. So first let's find $\sigma_V$.

From page 43 of the reader (just before equation 2.67) we find how we van find $\sigma_V$ if $\underline{V}$ is the product of two random variables. (Note that $\rho$ in this equation denotes the correlation coefficient, which is zero, since the measurements are independent.) We can easily guess what this equation will look like if $\underline{V}$ is the product of three random variables. We thus get

$$\sigma_V^2 = \sigma_{x_1}^2 \left(\overline{x}_2\overline{x}_3\right)^2 + \sigma_{x_2}^2 \left(\overline{x}_3\overline{x}_1\right)^2 + \sigma_{x_3}^2 \left(\overline{x}_1\overline{x}_2\right)^2 = 0.1^2 \left(6^2 + 20^2 + 30^2\right) = 13.36cm^6. \tag{7}$$

It follows that $\sigma_V = 3.66cm^3$ and thus that $\sigma_M = \rho\sigma_V = 8.9 \cdot 3.66 = 32.53g$.

## Question 7

Let $\underline{x}$ be a continuous random variable with a standard normal distribution. We consider the transformation $\underline{y} = |\underline{x}|$. The probability density function (PDF) of $\underline{y}$ is given as...

## Solution

We know that $\underline{y} \geq 0$. So for $y < 0$ the probability $P(y \leq \underline{y} \leq y + dy) = 0$. So $f_{\underline{y}}$ must be zero for $y < 0$.

But what if $y \geq 0$? In this case (for some value $y$) the event $\underline{y} = |\underline{x}| = y$ is twice as likely as the event $\underline{x} = y$. (We can rewrite the first case as $\underline{x} = y \cup -\underline{x} = y$. Both these events are equally likely, so thus $P(|\underline{x}| = y) = 2P(\underline{x} = y)$.) It follows that $f_{\underline{y}}(y)$ is also twice as large as $f_{\underline{x}}(y)$. In an equation this becomes

$$f_{\underline{y}}(y) = 2f_{\underline{x}}(y) = \frac{2}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} = \sqrt{\frac{2}{\pi}} e^{-\frac{y^2}{2}}. \tag{8}$$

## Question 8

Let $\underline{x}$ and $\underline{y}$ be random variables with standard deviation $\sigma_x$ and $\sigma_y$, respectively. The correlation coefficient of $\underline{x}$ and $\underline{y}$ is $\rho \neq 0$. The variance of $\underline{z} = \frac{1}{2}\underline{x} - \underline{y}$ equals...

## Solution

Let's define $\underline{v} = \frac{1}{2}\underline{x}$ and $\underline{w} = -\underline{y}$. It now follows that $\sigma_v = \frac{1}{2}\sigma_x$ and $\sigma_w = \sigma_y$. Also, $\underline{z} = \underline{v} + \underline{w}$. We now have

$$\sigma_z^2 = \sigma_v^2 + \sigma_w^2 + 2\rho_{vw}\sigma_v\sigma_w = \frac{1}{4}\sigma_x^2 + \sigma_y^2 + 2\rho_{vw}\sigma_v\sigma_w. \tag{9}$$

However, we don't know $\rho_{vw}$. But we do know $\rho_{xy}$. So we need to express $\rho_{vw}$ in $\rho_{xy}$. We know that the correlation coefficient between $\underline{x}$ and $\underline{y}$ is given by

$$C(\underline{x}, \underline{y}) = E(\underline{xy}) = \rho_{xy}\sigma_x\sigma_y. \tag{10}$$

Similarly, the correlation coefficient between $\underline{v}$ and $\underline{w}$ is given as

$$\rho_{vw}\sigma_v\sigma_w = C\left(\underline{v}, \underline{w}\right) = E\left(\underline{vw}\right) = E\left(-\frac{1}{2}\underline{xy}\right) = -\frac{1}{2}\rho_{xy}\sigma_x\sigma_y. \tag{11}$$

(We have silently assumed here that $\overline{v} = \overline{w} = 0$. This is not necessarily the case. However, it doesn't matter if $\overline{v} \neq 0$ or $\overline{w} \neq 0$. Although the computation is a bit more difficult, the result is the same.) Inserting the above finding into our previous equation will give us

$$\sigma_z^2 = \frac{1}{4}\sigma_x^2 + \sigma_y^2 - \rho\sigma_x\sigma_y. \tag{12}$$

(Note that we have simply written $\rho_{xy}$ and $\rho$.)

## Question 9

Let $\underline{x}_1$, $\underline{x}_2$ and $\underline{x}_3$ be independent random variables, all having a standard normal PDF. These random variables are combined into the three-dimensional random vector $\underline{x} = [\underline{x}_1, \underline{x}_2, \underline{x}_3]^T$. We define a new random vector

$$\underline{y} = \left[\underline{y}_1, \underline{y}_2, \underline{y}_3\right]^T = [2\underline{x}_1 - \underline{x}_2, \underline{x}_1 - \underline{x}_2 - \underline{x}_3, \underline{x}_1 + 3\underline{x}_3]^T . \tag{13}$$

Which components of $\underline{y}$ are positively correlated?

## Solution

We first note that all the random variables involved in this question have as mean zero. That simplifies things a bit.

If a pair of random variables is positively correlated, then $\rho > 0$. It follows that the covariance $C$ must also be bigger than zero. So let's try to find this covariance $C$ for all the pairs.

First let's examine the pair $\underline{y}_1$ and $\underline{y}_2$. We know that

$$C(\underline{y}_1, \underline{y}_2) = E\left((2\underline{x}_1 - \underline{x}_2)(\underline{x}_1 - \underline{x}_2 - \underline{x}_3)\right). \tag{14}$$

We can work out brackets for the above equation. We can also use the fact that $E$ is a linear operator. We then get

$$C(\underline{y}_1, \underline{y}_2) = 2E\left(\underline{x}_1^2\right) - 3E\left(\underline{x}_1\underline{x}_2\right) - 2E\left(\underline{x}_1\underline{x}_3\right) + E\left(\underline{x}_2^2\right) + E\left(\underline{x}_2\underline{x}_3\right). \tag{15}$$

Since $\underline{x}_1$ and $\underline{x}_2$ are independent, we know that $E(\underline{x}_1\underline{x}_2) = E(\underline{x}_1)E(\underline{x}_2) = 0 \cdot 0 = 0$. The same goes for the other pairs of random variables. So we find that

$$C(\underline{y}_1, \underline{y}_2) = 2E\left(\underline{x}_1^2\right) + E\left(\underline{x}_2^2\right) = 2\sigma_{x_1}^2 + \sigma_{x_2}^2 = 3\sigma^2. \tag{16}$$

In the last step we have used the fact that the random variables have the same standard deviation $\sigma$. Since $\sigma^2$ is obviously positive, we know that the covariance is also positive. So the random variables $\underline{y}_1$ and $\underline{y}_2$ are positively correlated.

We can do the same for $\underline{y}_1$ and $\underline{y}_3$. We then get

$$C(\underline{y}_1, \underline{y}_3) = 2\sigma_1^2 = 2\sigma^2. \tag{17}$$

This is evidently also positive. So $\underline{y}_1$ and $\underline{y}_3$ are positively correlated.

Now let's investigate the pair $\underline{y}_2$ and $\underline{y}_3$. We now find that

$$C(\underline{y}_2, \underline{y}_3) = \sigma_1^2 - 3\sigma_3^2 = -2\sigma^2. \tag{18}$$

This is evidently a negative number. So $\underline{y}_2$ and $\underline{y}_3$ are not positively correlated.

## Question 10

We consider the so-called Central Limit Theorem of Lindeberg-Levy. Let $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ be independent random variables, all having the same distribution with mean $\mu$ and variance $\sigma^2$. Then the probability density function of the random variable

$$\underline{z} = \frac{\left(\frac{1}{n}\sum_{i=1}^{n} \underline{x}_i\right) - \mu}{\sigma/\sqrt{n}} \tag{19}$$

converges to the standard normal density function for $n \to \infty$, i.e.,

$$f_{\underline{z}}(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \text{for} n \to \infty. \tag{20}$$

Then at the same time, for $n \to \infty$, the random variable $\underline{y} = \sum_{i=1}^{n} \underline{x}_i$ is normally distributed with mean $n\mu$ and standard deviation $\sigma\sqrt{n}$. Practically, $n \to \infty$ means $n > 30$.

As an application of this important theorem of probability theory, we consider the following example: A person sets out the track for a 100 meter running match by taking 100 steps. The length of each step is uniformly distributed on the interval (0.9 meter, 1.1 meter). Assume the length of the steps to be independent. The probability that the length of the track differs less than 1 meter from the required 100 meter is...

## Solution

The average length of the track that was set out is 100 meters. The standard deviation of one step is (according to table 2.2 (page 31) of the reader) equal to $\sigma_s = \sqrt{0.2^2/12} = 0.0577m$. This makes the standard deviation of the entire track $\sigma_t = \sigma_s\sqrt{100} = 0.577m$.

The probability that the length of the track is between 99 meter and 101 meter now can be found using the Central Limit Theorem. We assume it is normally distributed with the just found mean and standard deviation. We insert in our TI calculator the calculation $normalcdf(99, 101, 100, 0.577)$. The result is 0.9167.

## Question 11

Let $\underline{x}_1$ and $\underline{x}_2$ be independent random variables, both normally distributed. The mean of $\underline{x}_1$ and $\underline{x}_2$ are 2 and 5, respectively, whereas the variances $\underline{x}_1$ and $\underline{x}_2$ are 5 and 9, respectively. We define $\underline{y} = 3\underline{x}_1 - 2\underline{x}_2 + 1$. The probability $P(\underline{y} \leq 6)$ reads...

## Solution

We know that $\underline{x}_1$ and $\underline{x}_2$ are normally distributed. Also, all linear combinations of normal distributions are normal distributions. So $\underline{y}$ is also normally distributed. We only need to find its mean and its average.

The mean of $\underline{y}$ can be found using

$$\overline{y} = 3\overline{x}_1 - 2\overline{x}_2 + 1 = -3. \tag{21}$$

The standard deviation of $3\underline{x}_1$ is $3\sqrt{5}$. The standard deviation of $2\underline{x}_2$ is $2\sqrt{9} = 6$. The standard deviation of the number 1 is, of course, simply 0. So the standard deviation of $\underline{y}$ can now be found using

$$\sigma_y^2 = \left(3\sqrt{5}\right)^2 + 6^2 = 81. \tag{22}$$

It follows that $\sigma_y = \sqrt{81} = 9$. We can now find $P(\underline{y} \leq 6)$ using $normalcdf(-1E99, 6, -3, 9)$. The result is 0.8413.

## Question 12

Let $\underline{x}$ be an exponentially distributed random variable with parameter $\lambda = \frac{1}{2}$, i.e. the probability density function (PDF) of $\underline{x}$ is

$$f_{\underline{x}}(x) = \frac{1}{2} \exp\left(-\frac{x}{2}\right) \quad \text{for} \quad x \geq 0 \quad \text{and} \quad f_{\underline{x}}(x) = 0 \quad \text{otherwise}. \tag{23}$$

The conditional probability $P(\underline{x} > 3 | 1 < \underline{x} < 4)$ equals...

## Solution

Let's apply the definition of conditional probability. We find that

$$P(\underline{x} > 3 | 1 < \underline{x} < 4) = \frac{P(\underline{x} > 3 \cap 1 < \underline{x} < 4)}{P(1 < \underline{x} < 4)} = \frac{P(3 < \underline{x} < 4)}{P(1 < \underline{x} < 4)}. \tag{24}$$

To find this probability, we need to integrate the PDF. It follows that

$$P(\underline{x} > 3 | 1 < \underline{x} < 4) = \frac{\int_3^4 e^{-\frac{x}{2}}}{\int_1^4 e^{-\frac{x}{2}}} = \frac{e^{-3/2} - e^{-4/2}}{e^{-1/2} - e^{-4/2}} = 0.186. \tag{25}$$

## Question 13

An object is moving along a straight line. The following measurements $y_i$ of the object's position have been made at corresponding times $t_i$:

| $i$ | time $t_i$ (in $s$) | position $y_i$ (in m) |
|---|---|---|
| 1 | −1 | −2 |
| 2 | 0 | 0 |
| 3 | 1 | 3 |
| 4 | 2 | 5 |

We fit the data by a linear model $y = x_0 + vt$ using the method of least-squares curve fitting. The (unweighted) least-squares solution for the speed $v$ is...

## Solution

The least-squares solution $\hat{\mathbf{x}}$ can be found using

$$\hat{\mathbf{x}} = \left(A^T A\right)^{-1} A^T \mathbf{y}, \tag{26}$$

where we have

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \qquad \text{and} \qquad \mathbf{y} = \begin{bmatrix} -2 \\ 0 \\ 3 \\ 5 \end{bmatrix}. \tag{27}$$

We can first find that

$$\left(A^T A\right)^{-1} = \begin{bmatrix} 0.3 & -0.1 \\ -0.1 & 0.2 \end{bmatrix}. \tag{28}$$

It follows that

$$\hat{\mathbf{x}} = \begin{bmatrix} x_0 \\ v \end{bmatrix} = \begin{bmatrix} 0.3 \\ 2.4 \end{bmatrix}. \tag{29}$$

So we have $v = 2.4$.

## Question 14

The length $x$ of an object is measured three times with three different instruments, denoted $A$, $B$ and $C$. The three measurements and their precision (standard deviation) are list in the table below.

| instrument | measurement (cm) | precision (cm) |
|---|---|---|
| A | 5.19 | 0.4 |
| B | 5.27 | 0.5 |
| C | 5.21 | 0.2 |

All three measurements are uncorrelated. The Best Linear Unbiased Estimate (BLUE) of $x$, as a weighted average of the three measurements, and its standard deviation are...

## Solution

The variance matrix of the measurements is given as

$$Q_{yy} = \begin{bmatrix} 0.4^2 & 0 & 0 \\ 0 & 0.5^2 & 0 \\ 0 & 0 & 0.2^2 \end{bmatrix}. \tag{30}$$

Also, we are trying to solve the system $\mathbf{y} = A\mathbf{x}$, where $A = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$. The BLUE is now given by

$$\hat{x} = \left(A^T Q_{yy}^{-1} A\right)^{-1} A^T Q_{yy}^{-1} \mathbf{y}. \tag{31}$$

We can find that $A^T Q_{yy}^{-1} A = 141/4 = 35.25$ and also that $\left(A^T Q_{yy}^{-1} A\right)^{-1} = 4/141$. We eventually find that

$$\hat{x} = \frac{25y_1 + 16y_2 + 100y_3}{141}. \tag{32}$$

Inserting values gives $\hat{x} = 5.213$. Also, we can find that

$$\sigma_x^2 = \frac{25^2\sigma_{y_1}^2 + 16^2\sigma_{y_2}^2 + 100^2\sigma_{y_3}^2}{141^2}. \tag{33}$$

Inserting values, we find that $\sigma_x^2 = 4/141$. It follows that $\sigma_x = \sqrt{4/141} = 0.1684$.

## Question 15

Given is the linear model $E(\underline{y}) = Ax$, $D(\underline{y}) = \sigma^2 I_3$, with $x = [x_1, x_2]^T$ and

$$A = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}. \tag{34}$$

The standard deviation of the BLUE of $x_1$ is given as...

## Solution

The BLUE is, as always, given by

$$\hat{\mathbf{x}} = \left(A^T Q_{yy}^{-1} A\right)^{-1} A^T Q_{yy}^{-1} \mathbf{y}. \tag{35}$$

We can find that

$$\left(A^T Q_{yy}^{-1} A\right)^{-1} = \sigma^2 \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{bmatrix}. \tag{36}$$

It follows that

$$\hat{\mathbf{x}} = \begin{bmatrix} 2/3 & -1/3 & -1/3 \\ 1/3 & 1/3 & -2/3 \end{bmatrix} \mathbf{y}. \tag{37}$$

So we have

$$x_1 = \frac{2}{3}y_1 - \frac{1}{3}y_2 - \frac{1}{3}y_3. \tag{38}$$

The variance of $x_1$ is thus given by

$$\sigma_{x_1}^2 = \left(\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2\right)\sigma^2 = \frac{2}{3}\sigma^2. \tag{39}$$

It follows that the standard deviation is $\sigma_{x_1} = \sqrt{\frac{2}{3}}\sigma = 0.816\sigma$.

## Question 16

Given is the linear model $E(\underline{y}) = Ax$, $D(\underline{y}) = Q_{yy}$, where $A$ is an $m \times n$ matrix with rank$(A) = n$. Let $\hat{\underline{x}}$ be the weighted least-squares estimator (WLSE) of $x$, with weight matrix $W = \sigma^2 I_m$. The variance matrix of $\hat{\underline{z}} = F^T \hat{\underline{x}} + f_0$ is then given as...

## Solution

We know that $\underline{z} = F^T \hat{\underline{x}} + f_0$, with $F$ a constant matrix and $f_0$ a constant vector. Adding up a constant vector to a random vector doesn't change its variance matrix. So we may ignore $f_0$ when calculating the variance matrix. But what about $\hat{\underline{x}}$? Well, we know that

$$\hat{\underline{x}} = \left(A^T W A\right)^{-1} A^T W \underline{y}. \tag{40}$$

It follows that

$$\underline{z} = F^T \left(A^T W A\right)^{-1} A^T W \underline{y} + f_0. \tag{41}$$

We can now apply the variance propagation law from the reader (page 64, equation 2.120). We find that

$$Q_{zz} = F^T \left(A^T W A\right)^{-1} A^T W Q_{yy} W^T A \left(A^T W^T A\right)^{-1} F. \tag{42}$$

However, we know that $W = \sigma^2 I_m$. Inserting this into the above equation gives

$$Q_{zz} = F^T \left(A^T \sigma^2 I_m A\right)^{-1} A^T \sigma^2 I_m Q_{yy} \sigma^2 I_m A \left(A^T \sigma^2 I_m A\right)^{-1} F. \tag{43}$$

From linear algebra we know that $(cB)^{-1} = \frac{1}{c} B^{-1}$ (with $c$ a scalar and $B$ a matrix). Using this rule, we can reduce the above equation to

$$Q_{zz} = F^T \left(A^T A\right)^{-1} A^T Q_{yy} A \left(A^T A\right)^{-1} F. \tag{44}$$

## Question 17

The amplitude (signal-strength) of a received radar signal is measured once; observation $y$. The observable has a Rayleigh distribution with unknown parameter $x$:

$$f_{\underline{y}}(y|x) = \frac{y}{x^2} e^{-\frac{y^2}{2x^2}} \quad \text{with} \quad y \geq 0, x \geq 0. \tag{45}$$

Determine the Maximum Likelihood Estimate for the parameter $x$.

## Solution

To find the MLE, we need to find the $x$ which maximizes the PDF. In other words, its derivative (with respect to $x$) should be zero. This gives us the equation

$$-\frac{2y}{x^3} e^{-\frac{y^2}{2x^2}} + \frac{2y^2}{2x^3} \frac{y}{x^2} e^{-\frac{y^2}{2x^2}} = 0. \tag{46}$$

It follows that

$$\frac{y^3}{x^5} = 2\frac{y}{x^3}. \tag{47}$$

We know that $x > 0$, so we may rewrite this as $y^3 = 2yx^2$. If $y = 0$ all $x$ are valid solutions. So we assume that $y \neq 0$. We can then rewrite our equation to

$$x = \sqrt{\frac{1}{2} y^2} = \frac{1}{2}\sqrt{2} y. \tag{48}$$

Since this is the MLE, we denote it as $\hat{x}$. So $\hat{x} = \frac{1}{2}\sqrt{2}y$.

## Question 18

A vehicle is moving at constant speed, along a straight line. It started at $t = 0$ at $x = 0$. At time $t = 1, 2, 3$ the position of the vehicle is observed. The observation values are $y(t = 1) = 4$, $y(t = 2) = 11$ and $y(t = 3) = 16$. The observables all have standard deviation $\sigma = 1$ and are uncorrelated. Compute the BLUE for the velocity of the vehicle.

## Solution

We know that $x_0 = 0$, so our equation is $y = vt$. We have

$$y = \begin{bmatrix} 4 \\ 11 \\ 16 \end{bmatrix} \qquad \text{and} \qquad A = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}. \tag{49}$$

We can still find the BLUE using

$$\hat{\mathbf{x}} = \left(A^T Q_{yy}^{-1} A\right)^{-1} A^T Q_{yy}^{-1} \mathbf{y}. \tag{50}$$

Eventually we find that

$$v = \frac{1 \cdot 4 + 2 \cdot 11 + 3 \cdot 16}{14} = 5.286 = 74/14 = 5 + 2/7. \tag{51}$$

## Question 19

For the problem of Question 18, compute the standard deviation of the estimator for the position at $t = 4$.

## Solution

We know from the previous problem that our BLUE is

$$\underline{\hat{v}} = \frac{\underline{y}_1 + 2\underline{y}_2 + 3\underline{y}_3}{14}. \tag{52}$$

It follows that the variance of $\underline{\hat{v}}$ is given by

$$\sigma_{\hat{v}}^2 = \frac{\sigma^2 + 2^2\sigma^2 + 3^2\sigma^2}{14^2} = \frac{1}{14}\sigma^2 = \frac{1}{14}. \tag{53}$$

It follows that $\sigma_{\hat{v}} = 1/\sqrt{14}$. However, $\hat{v}$ denotes the initial velocity. We know that the initial position is still simply $x_0 = 0$. So the position $\underline{x}$ at $t = 4$ can be found using $\underline{x} = \underline{\hat{v}}t = 4\underline{\hat{v}}$. The standard deviation of this position now becomes $\sigma_x = 4\sigma_{\hat{v}} = 4/\sqrt{14} = \frac{2}{7}\sqrt{14}$.

## Question 20

Suppose that the standard deviation of the estimator $\underline{\hat{x}}$ is $\sigma_{\hat{x}} = 4$. The estimator is normally distributed. The 97.5% confidence region $[\hat{x} - \epsilon, \hat{x} + \epsilon]$ for $x$, centered at the estimate $\hat{x}$, extends to both sides by $\epsilon$. The one-sided length of the interval is...

## Solution

With the "one-sided length" it is meant that we should find the length of the interval from the average to one edge. In other words, we should just find $\epsilon$. To do this, we use a TI calculator. We insert the equation $normalcdf(-X, X, 0, 4) = 0.975$. The result is 8.9656.

(In the above computation have assumed that the average is 0. However, we could also assume it to be some value $A$. We then need to insert the equation $normalcdf(A - X, A + X, A, 4) = 0.975$. The result will, however, be exactly the same for every $A$.)

## Question 21

With a laser-disto-meter the same, unknown, distance is measured four times. The observables are normally distributed, uncorrelated and the measurement error standard deviation is 2 millimeter, as specified by the manufacturer. The observations are $y_1 = 5.241m$, $y_2 = 5.239m$, $y_3 = 5.236m$ and $y_4 = 5.244m$. Determine the squared weighted norm of the BLUE residual vector, and check the value of the quadratic form against its nominal distribution at the 1% significance level. The value of the quadratic form and the critical value read...

## Solution

We first need to find the BLUE of the measurements. It can be shown that this is simply the average value, so $\hat{x} = 5.240$. The BLUE residual vector now is $\hat{e} = \begin{bmatrix} 0.001 & -0.001 & -0.004 & 0.004 \end{bmatrix}$. But how do we find the squared norm? For the weighted least squares method (according to page 102 of the reader (just after equation 3.13)) we find that the squared (weighted) norm is $\hat{e}^T W \hat{e}$. In the BLUE method, the weight matrix $W$ corresponds to the inverse of the variance matrix, so we can say that $W = Q_{yy}^{-1}$. Since $Q_{yy} = \sigma^2 I_4$ we have $W = \left(\frac{1}{\sigma^2}\right) I_4$. It follows that

$$\hat{e}^T W \hat{e} = \frac{e_1^2 + e_2^2 + e_3^2 + e_4^2}{\sigma^2} = \frac{1 + 1 + 16 + 16}{4} = 8.5mm. \tag{54}$$

Now let's try to find the critical value of this squared norm. We have seen that, when finding this squared norm, we added up 4 squared values. In fact, we have added up 4 squared normally distributed random variables. And when we add up normally distributed random variables, we get a Chi-square distribution.

What amount of degrees of freedom does this distribution have? Well, we have 4 measurements, while we only have 1 unknown. So the distribution has $4 - 1 = 3$ degrees of freedom. Also we can assume its noncentrality parameter $\lambda$ is zero. So we are dealing with a central Chi-square distribution with 3 degrees of freedom. We can now look up the critical value on page 338 of the reader. Inserting the level of significance $\alpha = 0.01$ and the degrees of freedom $n = 3$ will give us the critical value 11.3449.

## Question 22

It is assumed that the IQ-score for the population of students in Aerospace Engineering is distributed as $N(\mu, \sigma^2)$, with $\sigma = 10$. The TU Delft students' average of $\mu = 110$ is the null hypothesis $H_0$. To test the alternative hypothesis, $H_a : \mu = 120$, that Aerospace students are smarter, a random sample is taken: the IQ of 16 AE students is measured and the mean comes out as 114.5. Are they smarter? Use the Best test. What is the decision at the 10% level of significance, and what is the decision at the 5% level of significance? Reject the null hypothesis $H_0$ and conclude 'yes, they're smarter'? The decisions at 10% and 5% level of significance respectively read...

## Solution

If $k$ is the critical value, and $\underline{x}$ is the IQ of the students, then the level of significance $\alpha$ is

$$\alpha = P(\underline{x} > k | H_0). \tag{55}$$

So we assume $H_0$; we assume that the average IQ of the 16 students is $\mu = 110$. We also know that the standard deviation of the IQ of one student is $\sigma_x = 10$. The standard deviation of the average IQ of all the 16 students that were tested now becomes $\sigma_x / \sqrt{n} = \sigma_x / \sqrt{16} = 2.5$. Thus we assume that $\underline{x} \sim N(110, 2.5)$.

Suppose $\alpha = 0.1$. In this case we have $P(\underline{x} > k) = 0.1$. Inserting the equation $normalcdf(X, 1E99, 110, 2.5) = 0.1$ in a TI calculator, and letting it solve it, gives $k = 113.2$. Since our measured value is above the critical value, we have to reject $H_0$. The students ARE smarter.

If $\alpha = 0.05$, we can do the same calculation. We now find $k = 114.1$. We arrive at the same conclusion: reject $H_0$.

## Question 23

For test statistic $\underline{y}$ with normal distribution, two simple hypotheses are put forward

$$H_0 : \underline{y} \sim N(0, 4) \quad \text{versus} \quad H_a : \underline{y} \sim N(\nabla, 4). \tag{56}$$

If, with a right-sided critical region, the type I and type II error probabilities are respectively $\alpha = 0.025$ and $\beta = 0.2546$, what is the value of parameter $\nabla$?

## Solution

To solve this question, we need to take two steps. First we use $H_0$ and $\alpha$ to find the critical value $k$. Then we use $k$, together with $H_a$ and $\beta$, to find $\nabla$.

We know that

$$\alpha = P(\underline{y} > k | H_0). \tag{57}$$

So we assume $H_0$. We then find the $k$ for which $P(\underline{y} > k)$ is equal to $\alpha$. We can find $k$ with a TI calculator. We insert the equation $normalcdf(X, 1E99, 0, 2) = 0.025$. We find that $k = 3.92$.

Now we use the definition

$$\beta = P(\underline{y} < k | H_a). \tag{58}$$

This time we insert in our calculator $normalcdf(-1E99, 3.92, X, 2) = 0.2546$. We find that $\nabla = 5.24$.

## Question 24

The Probability Density Function (PDF) of observable $\underline{y}$ is given as

$$f_{\underline{y}}(y|x) = \frac{1}{x} e^{-\frac{y}{x}} \quad \text{with} \quad y \geq 0 \tag{59}$$

with two competing hypotheses concerning parameter $x$

$$H_1 : x = x_1 = 1 \qquad \text{and} \qquad H_2 : x = x_2 = 2. \tag{60}$$

The costs for incorrect decisions are $C(x = x_1, \delta = x_2) = 2$ and $C(x = x_2, \delta = x_1) = 1$. Occurrence of hypothesis $H_1$ is twice as likely as hypothesis $H_2$. One observation $y$ is made. The best decision rule for deciding between $H_1$ and $H_2$ reads: reject $H_1$ (and accept $H_2$) if...

## Solution

Let's sort our data. We know that $C_{12} = 2$ and $C_{21} = 1$. Also $P(x_1) = 2/3$ and $P(x_2) = 1/3$. Using equation 5.15 (page 221) from the reader, we find that we need to reject $H_1$ if

$$\frac{f_y(y|x_1)}{f_y(y|x_2)} < \frac{P(x_2)C_{21}}{P(x_1)C_{12}} = \frac{1 \cdot 1/3}{2 \cdot 2/3} = \frac{1}{4}. \tag{61}$$

It follows that

$$e^{-y} < \frac{1}{4}\frac{1}{2} e^{-\frac{y}{2}}. \tag{62}$$

When working with inequalities, special care should be payed to the smaller-than/greather-than sign. For example, if you multiply both sides of the equation by $-1$, the sign flips. So let's not do that.

We can multiply both sides of the equation by $8e^y$. (This quantity is positive for all $y$, so the smaller-than sign doesn't change.) We get

$$8 < e^{\frac{y}{2}}. \tag{63}$$

Since both sides of the equation are positive for all $y$, we can take the natural logarithm of both sides. The natural logarithm is an increasing function, so the smaller-than sign still doesn't change. This time we get

$$\ln 8 < \frac{y}{2}. \tag{64}$$

This reduces to

$$y > 2 \ln 8. \tag{65}$$

So we should reject $H_1$ if $y > 2 \ln 8$.