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...

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?

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?

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)?

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?

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?

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...

Question 8

Let $\underline{x}$ and $\underline{y}$ be random variables with standard deviation σ_x and σ_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...

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

$$\underline{y} = \left[\underline{y}_1, \underline{y}_2, \underline{y}_3\right]^T = [2\underline{x}_1 - \underline{x}_2, \underline{x}_1 - \underline{x}_2 - \underline{x}_3, \underline{x}_1 + 3\underline{x}_3]^T . \tag{1}$$

Which components of $\underline{y}$ are 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 μ and variance σ^2. Then the probability density function of the random variable

$$\underline{z} = \frac{\left(\frac{1}{n}\sum_{i=1}^{n}\underline{x}_i\right) - \mu}{\sigma/\sqrt{n}} \tag{2}$$

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

$$f_{\underline{z}}(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \text{for} n \to \infty. \tag{3}$$

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...

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...

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

$$f_{\underline{x}}(x) = \frac{1}{2}\exp\left(-\frac{x}{2}\right) \quad \text{for} \quad x \geq 0 \quad \text{and} \quad f_{\underline{x}}(x) = 0 \quad \text{otherwise.} \tag{4}$$

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

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...

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...

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

$$A = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix}. \tag{5}$$

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

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 $\text{rank}(A) = n$. Let $\underline{\hat{x}}$ be the weighted least-squares estimator (WLSE) of x, with weight matrix $W = \sigma^2 I_m$. The variance matrix of $\underline{\hat{z}} = F^T \underline{\hat{x}} + f_0$ is then given as...

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:

$$f_{\underline{y}}(y|x) = \frac{y}{x^2} e^{-\frac{y^2}{2x^2}} \quad \text{with} \quad y \geq 0, x \geq 0. \tag{6}$$

Determine the Maximum Likelihood Estimate for the parameter x.

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.

Question 19

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

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 ϵ. The one-sided length of the interval is...

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...

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...

Question 23

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

$$H_0 : \underline{y} \sim N(0, 4) \quad \text{versus} \quad H_a : \underline{y} \sim N(\nabla, 4). \tag{7}$$

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 ∇?

Question 24

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

$$f_{\underline{y}}(y|x) = \frac{1}{x} e^{-\frac{y}{x}} \quad \text{with} \quad y \geq 0 \tag{8}$$

with two competing hypotheses concerning parameter x

$$H_1 : x = x_1 = 1 \qquad \text{and} \qquad H_2 : x = x_2 = 2. \tag{9}$$

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...