Probability Theory
Exam July 2007 - Solutions

Question 1

Let A and B be events in a probability space. The probability of event A is $P(A) = 1/3$ and the conditional probability of B given A^c is $P(B|A^c) = 1/4$. The probability that A or B occurs, i.e. $P(A \cup B)$, equals...

Solution

We know that $P(A \cup B) = 1 - P(A^c \cap B^c)$. From the definition of probability, we also know that $P(A^c \cap B^c) = P(B^c|A^c)P(A^c)$. It now follows that

$$P(A \cup B) = 1 - P(A^c \cap B^c) = 1 - P(B^c|A^c)P(A^c) = 1 - (1 - P(B|A^c))(1 - P(A)) = 1/2. \tag{1}$$

Question 2

Let E and F be two events for which the probability that at least one of them occurs is 3/4. The probability that neither E nor F occurs is...

Solution

Given is that $P(E \cup F) = 3/4$. Asked is for $P(E^c \cap F^c)$. This is simply equal to

$$P(E^c \cap F^c) = 1 - P(E \cup F) = 1 - 3/4 = 1/4. \tag{2}$$

Question 3

One tosses a fair coin twice. The two events of interest are: $A = \{\text{first toss is a head}\}$ and $B = \{\text{second toss is a head}\}$. Are A and B independent? And are they disjoint?

Solution

The two tosses are, of course, independent. If you throw a head the first time, there will still be a 50% chance you throw a head the next time. They are, however, not disjoint. Both events contain the same element, being "head".

Question 4

On January 28, 1986 the space shuttle Challenger exploded about one minute after the launch. The cause of the disaster was explosion of the main fuel tank, caused by flames of hot gas erupting from one of the solid rocket boosters. These rocket boosters are manufactured in segments, joint together with O-rings. Each rocket booster has three O-rings and per launch two rocket boosters are used, so in total six O-rings each time. Based on data on the number of failed O-rings, available from previous launches, it was found that the probability p that an individual O-ring fails depends on the launch temperature t (in degrees Fahrenheit) according to

$$p = \frac{\exp(a + bt)}{1 + \exp(a + bt)}, \tag{3}$$

with $a = 5.085$ and $b = -0.1156$. Hence, p increases with decreasing launch temperature. At the time of the fatal launch of the Challenger, t was extremely low: 31 degrees Fahrenheit. Although the above formula is based on data for which $t > 50$ degrees Fahrenheit, let us use this formula also for $t = 31$ degrees Fahrenheit. Then, the probability of at least one O-ring failing during the 1986 Challenger launch equals...

Solution

By substituting values, we can find that $p = 0.8177$. Let's call $\underline{x}$ the number of O-rings failing. We can now see that the chance that at least one O-ring fails is

$$P(\underline{x} \geq 1) = 1 - P(\underline{x} = 0) = 1 - (1-p)^6 = 0.99996. \tag{4}$$

Question 5

A candidate pilot is declared suited for the job if his length lies in between certain boundaries. Now, the length of candidate pilots is normally distributed with mean 175 cm and standard deviation 8.5 cm. The goal is to declare 10% of the candidate pilots unfit based on their lengths, in such a way that the number of too short candidates equals the number of too long candidates. Which boundaries should one choose?

Solution

Let's denote the length of a candidate by $\underline{x}$. The average of the length is $\overline{x} = 175cm$. Also $\sigma_x = 8.5cm$. We know that 5% of the candidates should be too short and 5% should be too long. So we should define some variable r, such that $P(\underline{x} < \overline{x} - r) = 0.05$ and also $P(\underline{x} > \overline{x} - r) = 0.05$. In other words, the chance that some one is of the right length is $P(\overline{x} - r < \underline{x} < \overline{x} + r) = 0.9$.

Using the TI calculator, we can insert the equation $normalcdf(175 - X, 175 + X, 175, 8.5) = 0.9$. Letting the calculator solve for X will give $X = 13.98$. So anyone with his length in the range $[161, 189]$ will be approved for the length-criterion.

Question 6

A machine fastens plastic screw-on caps onto containers of motor oil. If the machine applies more torque than the cap can withstand, the cap will break. Both the applied torque and the strength of the caps vary. The capping machine torque is a normally distributed random variable with mean $0.79Nm$ and standard deviation $0.10Nm$. The cap strength, being the torque that would break the cap, is also a normally distributed random variable with mean $1.13Nm$ and standard deviation $0.14Nm$. Assume that the cap strength and the applied torque are independent. The probability that a cap will break while being fastened by the capping machine equals...

Solution

Let's call the machine torque $\underline{x}_1$ and the cap strength $\underline{x}_2$. Now we define $\underline{y} = \underline{x}_2 - \underline{x}_1$. We can see that if $\underline{y} < 0$, then the cap will break. So we need to find $P(\underline{y} < 0)$.

Since $\underline{x}_1$ and $\underline{x}_2$ are normally divided, also $\underline{y}$ is normally divided. Its mean is $\overline{y} = \overline{x}_2 - \overline{x}_1 = 1.13 - 0.79 = 0.34Nm$. Its variance is given by $\sigma_y^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 = 0.10^2 + 0.14^2 = 0.0296$. It follows that the standard deviation is $\sigma_y = 0.172$. Using the TI calculator, we can now find that $P(\underline{y} < 0) = normalcdf(1E - 99, 0, 0.34, 0.172) = 0.024$.

Question 7

Let $\underline{x}_1$ and $\underline{x}_2$ be random variables. The mean and standard deviation of $\underline{x}_1$ are 5.0 and 2.9, whereas the mean and standard deviation of $\underline{x}_2$ are 13.2 and 17.6. The correlation coefficient of $\underline{x}_1$ and $\underline{x}_2$ is -0.11. The standard deviation of the random variable $\underline{y} = 0.2\underline{x}_1 + 0.8\underline{x}_2$ equals...

Solution

We know that the correlation coefficient of $\underline{x}_1$ and $\underline{x}_2$ is -0.11. The covariance then becomes $C(\underline{x}_1, \underline{x}_2) = -0.11 \cdot 2.9 \cdot 17.6 = -5.6144$. Also note that the covariance of a random variable with itself is the variance, so $C(\underline{x}_1, \underline{x}_1) = \sigma_{x_1}^2$.

Now let's consider $\underline{y}' = 5\underline{y} = \underline{x}_1 + 4\underline{x}_2$. The variance of $\underline{y}'$ now can be found using the covariance. It follows that

$$\sigma_{y'}^2 = \sigma_{x_1}^2 + 2 \cdot 4 \cdot C(\underline{x}_1, \underline{x}_2) + 4^2 \cdot \sigma_{x_2}^2 = 2.9^2 - 8 \cdot 5.6144 + 16 \cdot 17.6^2 = 4919.65. \tag{5}$$

This means that the standard deviation of $\underline{y}'$ is $\sigma_{y'} = 70.14$. The standard deviation of $\underline{y}$ thus becomes $\sigma_y = \sigma_{y'}/5 = 14.03$.

Question 8

We want to determine the volume of a cylinder by means of measuring the cylinder length L and radius r once. Before the actual measurement, we would like to assess the precision with which the cylinder volume can be determined. We know that the cylinder length is approximately $10cm$ and that the cylinder radius is approximately $2cm$. Assuming that the length and radius measurement are independent and have a standard deviation of $1mm$, what is the standard deviation of the volume, to a first-order approximation?

Solution

The volume, the length and the radius are, in fact, random variables. So we have that $\underline{V} = \pi \underline{r}^2 \underline{L}$. Thus $\underline{V}$ is a function of $\underline{r}$ and $\underline{L}$. Since $\underline{r}$ and $\underline{L}$ are independent, we know that their correlation ρ is 0. From the reader, page 43 (just before equation 2.67), we can find that

$$\sigma_V^2 = \sigma_L^2 \left(\pi \overline{r}^2\right)^2 + \sigma_r^2 \left(2\pi \overline{r}\overline{L}\right)^2 = 0.1^2 \left(\pi \cdot 2^2\right)^2 + 0.1^2 \left(2\pi \cdot 2 \cdot 10\right)^2 = 159.5cm^6. \tag{6}$$

The standard deviation of the volume now is $\sigma_V = 12.63cm^3$.

Question 9

In a certain country it is established that 0.5% of the population suffers from a certain disease. For this disease there exists a test that gives the correct diagnosis for 80% of healthy persons and for 98% of sick persons. A person is tested and found sick. The probability that the diagnosis is wrong, i.e. that the person is actually healthy, equals...

Solution

This is just a basic application of Bayes' rule. The probability $P(\text{healthy}|\text{diag.sick})$ that someone who is diagnosed sick is actually healthy is

$$P(\text{healthy}|\text{diag. sick}) = \frac{P(\text{diag. sick}|\text{healthy})P(\text{healthy})}{P(\text{diag. sick}|\text{healthy})P(\text{healthy}) + P(\text{diag. sick}|\text{sick})P(\text{sick})} = \ldots \tag{7}$$

$$\ldots = \frac{20 \cdot 99.5}{20 \cdot 99.5 + 98 \cdot 0.5} = 0.976 = 97.6\%. \tag{8}$$

Question 10

Let $\underline{x}_1$ and $\underline{x}_2$ be independent and normally distributed random variables. The mean of both $\underline{x}_1$ and $\underline{x}_2$ is 0. The variance of both $\underline{x}_1$ and $\underline{x}_2$ equals 3. Consider the linear transformation

$$\underline{y}_1 = \underline{x}_1 + \underline{x}_2 \qquad \text{and} \qquad \underline{y}_2 = \underline{x}_1 - 2\underline{x}_2. \tag{9}$$

Then the joint PDF of $\underline{y}_1$ and $\underline{y}_2$ reads...

Solution

We first will determine the matrix Q_{yy}. We can find that $\sigma_{y_1}^2 = 6$ and $\sigma_{y_2}^2 = 15$. We also have as correlation

$$C(\underline{y}_1, \underline{y}_2) = E\left((\underline{y}_1 - \overline{y}_1)(\underline{y}_2 - \overline{y}_2)\right) = E(\underline{y}_1 \underline{y}_2), \tag{10}$$

where we have used that $\overline{y}_1 = \overline{y}_2 = 0$. It follows that

$$E(\underline{y}_1 \underline{y}_2) = E\left((\underline{x}_1 + \underline{x}_2)(\underline{x}_1 - 2\underline{x}_2)\right) = E\left(\underline{x}_1^2 - \underline{x}_1 \underline{x}_2 - 2\underline{x}_2^2\right) = E(\underline{x}_1^2) - E(\underline{x}_1 \underline{x}_2) - 2E(\underline{x}_2^2). \tag{11}$$

We know that $\underline{x}_1$ and $\underline{x}_2$ are independent, so $E(\underline{x}_1 \underline{x}_2) = E(\underline{x}_1)E(\underline{x}_2) = 0$. We also know that $E(\underline{x}_1) = E(\underline{x}_2) = \sigma^2$. It follows that the correlation between $\underline{y}_1$ and $\underline{y}_2$ is $C(\underline{y}_1, \underline{y}_2) = -\sigma^2 = -3$. This gives us the matrix Q_{yy}, being

$$Q_{yy} = \begin{bmatrix} 6 & -3 \\ -3 & 15 \end{bmatrix} \tag{12}$$

Since $\underline{y}_1$ and $\underline{y}_2$ are linear functions of normal distributions, they theirselves are normal distributions. Together they thus form a multivariate normal distribution. The equation for this is

$$f_{\underline{y}_1, \underline{y}_2}(y_1, y_2) = \frac{1}{\sqrt{\det(2\pi Q_{yy})}} e^{\left(-\frac{1}{2}(\mathbf{y} - \overline{\mathbf{y}})^T Q_{yy}^{-1} (\mathbf{y} - \overline{\mathbf{y}})\right)}. \tag{13}$$

We can find that $\sqrt{\det(2\pi Q_{yy})} = 18\pi$. We can also find that

$$Q_{yy}^{-1} = \begin{bmatrix} 5/27 & 1/27 \\ 1/27 & 2/27 \end{bmatrix} \tag{14}$$

It follows that

$$-\frac{1}{2}(\mathbf{y} - \overline{\mathbf{y}})^T Q_{yy}^{-1} (\mathbf{y} - \overline{\mathbf{y}}) = -\frac{5y_1^2 + 2y_1 y_2 + 2y_2^2}{54}. \tag{15}$$

Now we know what our multivariate normal distribution will be. It is

$$\frac{1}{18\pi} e^{\left(-\frac{5y_1^2 + 2y_1 y_2 + 2y_2^2}{54}\right)}. \tag{16}$$

Question 11

Let $\underline{x}$ be an exponentially distributed random variable with parameter $\lambda = 1/5$. The conditional probability $P(x < 5 | 3 < x < 6)$ equals...

Solution

We simply apply the definition of probability. We then get

$$P(x < 5 | 3 < x < 6) = \frac{P(x < 5 \cap 3 < x < 6)}{P(3 < x < 6)} = \frac{P(3 < x < 5)}{P(3 < x < 6)}. \tag{17}$$

For an exponential distribution, we have (for $x \geq 0$) as PDf

$$f_{\underline{x}}(x) = \lambda e^{-\lambda x}. \tag{18}$$

To find the probability, we need to integrate this PDF. We then get

$$P(x < 5 | 3 < x < 6) = \frac{\int_3^5 \lambda e^{-\lambda x}}{\int_3^6 \lambda e^{-\lambda x}} = \frac{\left[-e^{-\lambda x}\right]_3^5}{\left[-e^{-\lambda x}\right]_3^6} = \frac{e^{-3/5} - e^{-5/5}}{e^{-3/5} - e^{-6/5}} = 0.731. \tag{19}$$

Question 12

The random variable $\underline{x}$ is uniformly distributed on the interval $(0, 1)$. Then the PDF of the random variable $\underline{y} = -\ln \underline{x}$ reads...

Solution

We know that the PDF of $\underline{x}$ is given by

$$f_{\underline{x}}(x) = \begin{cases} 1 & \text{for } 0 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

We also know that $\underline{y} = g(\underline{x})$, where $g(x) = -\ln x$. It follows that the inverse of $g(x)$ is $g^{-1}(y) = e^{-y}$. We now apply the rule, stating that

$$f_{\underline{y}}(y) = f_{\underline{x}}(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right| = f_{\underline{x}}(e^{-y})| - e^{-y}|. \tag{21}$$

We first note that e^{-y} is also positive, so $| - e^{-y}| = e^{-y}$. We also see that $f_{\underline{x}}(x) = 1$ if $0 \leq x \leq 1$, so if $y \geq 0$. It is zero otherwise. It follows that

$$f_{\underline{y}}(y) = \begin{cases} e^{-y} & \text{for } y \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

Question 13

The random variable $\underline{w}$ is uniformly distributed on the interval $(\pi, 2\pi)$. What can we say about $E(\sin(\underline{w}))$ and $\sin(E(\underline{w}))$? Which one is bigger? And if they are equal, are they also equal to zero?

Solution

Since $\underline{w}$ is uniformly distributed over $(\pi, 2\pi)$, we know that $E(\underline{w})$ is in the middle of that interval, so $E(\underline{w}) = \frac{3}{2}\pi$. It follows that $\sin(E(\underline{w})) = \sin((3/2)\pi) = -1$.

Now let's look at $E(\sin(\underline{w}))$. Since we know that $\underline{w}$ is in the interval $(\pi, 2\pi)$, we know that $\sin(\underline{w})$ is between 0 and -1. The mean of $\sin(\underline{w})$ is thus also somewhere between 0 and -1. Although we don't know its value, we can already conclude that $E(\sin(\underline{w})) > \sin(E(\underline{w}))$.

Question 14

Given is the linear model $E(\underline{y}) = Ax$, $D(\underline{y}) = 4I_m$ with the $m \times 1$ matrix $A = [1, \ldots, 1]^T$. $\underline{\hat{x}}$ is the Best Linear Unbiased Estimator (BLUE) of x. There is a requirement that the variance of $\underline{\hat{x}}$ is at most equal to 0.5. How many observations m should one then at least take?

Solution

We use the Guass-Markov theorem, which states that

$$\underline{\hat{\mathbf{x}}} = \left(A^T Q_{yy}^{-1} A\right)^{-1} A^T Q_{yy}^{-1} \underline{\mathbf{y}}. \tag{23}$$

It can be shown that $A^T Q_{yy}^{-1} A = 4/m$, so also $\left(A^T Q_{yy}^{-1} A\right)^{-1} = m/4$. It now follows that

$$\underline{\hat{\mathbf{x}}} = \begin{bmatrix} 1/m & 1/m & \ldots & 1/m \end{bmatrix} \underline{\mathbf{y}}. \tag{24}$$

Let's take a look at the above equation. We know we have taken m measurements $y_1, y_2, \ldots, y_m$. And the estimator $\underline{\hat{x}}$ just takes the average of all those measurements! We also know that one measurement has variance $\sigma_y^2 = 4$ (since if $m = 1$, then $\sigma_y^2 = D(\underline{y}) = 4I_1 = 4$). It follows that the variance of $\underline{\hat{x}}$ is $\sigma_{\hat{x}}^2 = \sigma_y^2/m = 4/m$. Since we want to have $\sigma_{\hat{x}}^2 = 0.5$, it follows that $m = 8$.

Question 15

Three points A, B and C lie on a straight line. All pairwise distances $y_1 = AB$, $y_2 = BC$ and $y_3 = AC$ are measured. Thus, the observables are $\underline{y}_i = y_i + \underline{e}_i$, whereby it may assumed that $E(\underline{e}_i) = 0$ and $D(\underline{e}_i) = 3cm^2$ for $i = 1, 2, 3$, and $C(\underline{e}_i, \underline{e}_j) = 0$ for $i \neq j$. The variance of the BLUE of y_3 is then given as...

Solution

This is a rather difficult question, since you need to come up with some insights of your own. We know we have three measurements y_1, y_2 and y_3 of the distances AB, BC and AC, respectively. Ask yourself, in what ways can we make an estimation x_3 of the distance AC?

We can make an estimation in two ways. First of all, we can say that $x_3 \approx y_3$. However, we can also say that $x_3 \approx y_1 + y_2$. So we have two ways to make the estimation x_3. This gives us the system of equations $\mathbf{y} = Ax_3$, with

$$\mathbf{y} = \begin{bmatrix} y_1 + y_2 \\ y_3 \end{bmatrix} \qquad \text{and} \qquad A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \tag{25}$$

We should find x_3 using the BLUE method. However, to apply that method, we need the variance matrix Q_{yy}. We know that the variance of $\underline{y}_3$ is simply $D(\underline{y}_3) = \sigma^2$. The variance of $\underline{y}_1 + \underline{y}_2$ is equal to $2\sigma^2$. (Note that in this step we used the fact that the covariance of those two random variables is zero.) Also, the covariance between $\underline{y}_1 + \underline{y}_2$ and $\underline{y}_3$ is zero. So the variance matrix finally becomes

$$Q_{yy} = \begin{bmatrix} 2\sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}. \tag{26}$$

We can now find the BLUE, being

$$x_3 = \left(A^T Q_{yy}^{-1} A\right)^{-1} A^T Q_{yy}^{-1} \mathbf{y} = \frac{1}{3}(y_1 + y_2) + \frac{2}{3} y_3. \tag{27}$$

So the variance of x_3 thus becomes

$$\sigma_{x_3}^2 = \left(\left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{2}{3}\right)^2 \right) \sigma^2 = \frac{2}{3}\sigma^2 = \frac{2}{3} \cdot 3 = 2cm^2. \tag{28}$$

Question 16

Given is the linear model $E(\underline{y}) = Ax$ and $D(\underline{y}) = \sigma^2 I_3$, where $y = [2, 1, 1/2]^T$, $x = [x_1, x_2]^T$ and

$$A = \begin{bmatrix} 1 & 0 \\ -1 & 2 \\ 2 & 1 \end{bmatrix} \tag{29}$$

Let $\underline{\hat{x}}_1$ denote the BLUE of x_1. We want the standard deviation of $\underline{\hat{x}}_1$ to be smaller than or at most equal to 1. What is the maximum allowable value of the standard deviation σ of the observations?

Solution

Once more we use the Guass-Markov theorem, which states that

$$\underline{\hat{\mathbf{x}}} = \left(A^T Q_{yy}^{-1} A\right)^{-1} A^T Q_{yy}^{-1} \underline{\mathbf{y}}. \tag{30}$$

We now have

$$\left(A^T Q_{yy}^{-1} A\right)^{-1} = \begin{bmatrix} 1/6 & 0 \\ 0 & 1/5 \end{bmatrix} \tag{31}$$

It follows that

$$\underline{\hat{\mathbf{x}}} = \begin{bmatrix} 1/6 & -1/6 & 1/3 \\ 0 & 2/5 & 1/5 \end{bmatrix} \mathbf{y}. \tag{32}$$

If we only look at $\underline{\hat{x}}_1$, then we find that

$$\underline{\hat{x}}_1 = \frac{1}{6}\left(\underline{y}_1 - \underline{y}_2 + 2\underline{y}_3\right). \tag{33}$$

The standard deviation of $\underline{y}_3$ is σ, so the standard deviation of $2\underline{y}_3$ is 2σ. Its variance is thus $4\sigma^2$. The variances of $\underline{y}_1$ and $-\underline{y}_2$ are both σ^2. So the variance of $\underline{y}_1 - \underline{y}_2 + 2\underline{y}_3$ becomes $\sigma^2 + \sigma^2 + 4\sigma^2 = 6\sigma^2$. The standard deviation is thus $\sqrt{6}\sigma$. Now we look at the standard deviation of $\frac{1}{6}\left(\underline{y}_1 - \underline{y}_2 + 2\underline{y}_3\right)$. This becomes $\sigma/\sqrt{6}$. If this should be at most 1, then σ should be at most $\sqrt{6}$.

Question 17

Consider the Gauss-Markov model $\underline{y} \sim N_m(Ax, Q_{yy})$. Are $\underline{\hat{x}}$ and $\underline{y}$ dependent? And are $\underline{\hat{x}}$ and $\underline{\hat{e}}$ dependent?

Solution

We know that $\underline{\hat{x}}$ is a linear function of $\underline{y}$, according to the Gauss-Markov theorem. So clearly they are dependent. However, according to the reader (page 144, just before equation 3.119), the random variables $\underline{\hat{x}}$ and $\underline{\hat{e}}$ are independent.

Question 18

We want to fit a plane $y = A + Bt + Cz$ to the four points $y = 3$ at $(t, z) = (1, 1)$, $y = 6$ at $(t, z) = (0, 3)$, $y = 5$ at $(t, z) = (2, 1)$ and $y = 0$ at $(t, z) = (0, 0)$. The system of normal equations from which the least squares solution for the unknown parameters can be obtained, is...

Solution

We recognize this as a problem $\mathbf{y} \approx S\mathbf{x}$, with

$$\mathbf{y} = \begin{bmatrix} 3 \\ 6 \\ 5 \\ 0 \end{bmatrix} \qquad S = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 3 \\ 1 & 2 & 1 \\ 1 & 0 & 0 \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} A \\ B \\ C \end{bmatrix} \tag{34}$$

We have used S as a matrix, because the variable A is already used. We know that the least squares solution can be found using $\hat{\mathbf{x}} = \left(S^T S\right)^{-1} S^T \mathbf{y}$, or, equivalently,

$$S^T S \hat{\mathbf{x}} = S^T \mathbf{y}. \tag{35}$$

By inserting values, and working things out, we get

$$\begin{bmatrix} 4 & 3 & 5 \\ 3 & 5 & 3 \\ 5 & 3 & 11 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 14 \\ 13 \\ 26 \end{bmatrix} \tag{36}$$

From this the least-squares solution can be obtained.

Question 19

At $t = 0$ ESA discovers an asteroid headed directly for Earth. In the two days ($t = 1$ and $t = 2$) after the discovery they measure the distance toward the object (d) using radar. By coincidence, they also find a useable observation taken three days before the discovery ($t = -3$) and determine the distance at that moment. The measured distances are 12.6, 9.1, and 8.3 ($\times 10^6 km$) for $t = -3$, $t = 1$, and $t = 2$, respectively. Assume that the asteroid moves directly toward us with an unknown constant velocity (v) and that the position $d(0)$ at $t = 0$ is also unknown. From the least squares solution for $d(0)$ and v, we estimate the time at which the asteroid will hit the Earth. This time equals...

Solution

This system can be described by $\mathbf{y} \approx A\mathbf{x}$, or, equivalently

$$\begin{bmatrix} 12.6 \\ 9.1 \\ 8.3 \end{bmatrix} \approx \begin{bmatrix} 1 & -3 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} d_0 \\ v_0 \end{bmatrix} \tag{37}$$

The least squares solutions is now given by $\hat{\mathbf{x}} = \left(A^T A\right)^{-1} A^T \mathbf{y}$. We can find that

$$\left(A^T A\right)^{-1} = \begin{bmatrix} 1/3 & 0 \\ 0 & 1/14 \end{bmatrix} \tag{38}$$

Working things out further, we finally get that $d_0 = 10$ and $v_0 = -0.864$. The minus sign is present, because the distance d is decreasing. The time until impact is thus $-d_0/v_0 = 11.57$ days.

Question 20

For testing observable $\underline{y}$ with normal distribution, two simple hypotheses are put forward: $H_0 : \underline{y} \sim N(0, 4)$ and $H_a : \underline{y} \sim N(-3, 4)$ If the type I error probability is $\alpha = 0.05$ using a left sided critical region, which of the following values is the power of the test?

Solution

The variance of both distributions is 4, so the standard deviation is 2. First let's find the border of the critical region. We will call this border k. We know that $P(\underline{y} > k|H_0) = 0.95$. To find k, we use a TI calculator, and let it solve the equation $normalcdf(X, 1E99, 0, 2) = 0.95$. We get as result $X = -3.29$. The power of the test γ is the probability $P(\underline{y} < k|H_a)$. We can find this using $normalcdf(-1E99, -3.29, -3, 2)$. We get as result 0.4424.

Question 21

We have four observations y_i with $i = 1, \ldots, 4$ to determine a single unknown parameter x, according to $E(y_i) = x$. The observations are $y_1 = 5.0$, $y_2 = 6.4$, $y_3 = 4.8$ and $y_4 = 7.0$. The observables, which are normally distributed, are uncorrelated and all have standard deviation $\sigma = 1.4$. Determine the squared norm of the BLUE residual vector, and check this squared norm against its nominal distribution at the 10% significance level.

The squared norm and the critical value respectively read...

Solution

We can find the BLUE of $\underline{x}$. It can be shown that this is in fact just the average of the measurements.

Taking the average of the four y values will give $x = 5.8$. The residual vector now becomes

$$\epsilon = \begin{bmatrix} 5.0 \\ 6.4 \\ 4.8 \\ 7.0 \end{bmatrix} - \begin{bmatrix} 5.8 \\ 5.8 \\ 5.8 \\ 5.8 \end{bmatrix} = \begin{bmatrix} -0.8 \\ 0.6 \\ -1.0 \\ 1.2 \end{bmatrix} \tag{39}$$

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 $W = Q_{yy}^{-1}$. It follows that

$$\hat{e}^T Q_{yy}^{-1} \hat{e} = \begin{bmatrix} -0.8 & 0.6 & -1.0 & 1.2 \end{bmatrix} \begin{bmatrix} 1/1.4 & 0 & 0 & 0 \\ 0 & 1/1.4 & 0 & 0 \\ 0 & 0 & 1/1.4 & 0 \\ 0 & 0 & 0 & 1/1.4 \end{bmatrix} \begin{bmatrix} -0.8 \\ 0.6 \\ -1.0 \\ 1.2 \end{bmatrix} = 1.7551. \tag{40}$$

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 its noncentrality parameter λ is zero. (Since $E(y_i) = x$ we have $E(\hat{e}_i) = 0$, which implies that $\bar{\hat{e}}_i^2 = 0$ and thus also $\lambda = 0$.) 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 $\alpha = 0.1$ and the degrees of freedom $n = 3$ will give us the critical value 6.2514.

Question 22

A random variable $\underline{x}$ is distributed as follows

$$\underline{x} \sim N(8, 9). \tag{41}$$

What is the probability that $x > 11$?

Solution

The variance is 9, so the standard deviation is 3. The asked probability can be found using a TI calculator, using $\mathit{normalcdf}(11, 1E99, 8, 3) = 0.1587$. That's all.

Question 23

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

$$f_y(y|x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(y-x)^2} \tag{42}$$

with two competing hypotheses concerning parameter x

$$H_1 : x = a_1 \qquad \text{and} \qquad H_2 : x = a_2 \tag{43}$$

with $a_2 > a_1$. The hypotheses are equally likely. 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

We can immediately see that $\underline{y}$ is a normally distributed random variable with mean x. We know that we should reject H_1 if

$$\frac{f_{\underline{y}}(y|H_1)}{f_{\underline{y}}(y|H_2)} < \frac{P(a_2)c_2}{P(a_1)c_1}. \tag{44}$$

We know that the two hypothesis are equally likely, so $P(a_1) = P(a_2)$. Also, nothing has been mentioned about costs, so we assume that they are equal too. It follows that $f_{\underline{y}}(y|H_1) = f_{\underline{y}}(y|H_2)$. We can solve this equation algebraically, but we can also solve this question in an easier way, using plain logic. $\underline{y}|H_1$ and $\underline{y}|H_2$ are both normal distributions with equal standard deviation. So the value of y at which their PDFs are equal is exactly between their means. Their means are a_1 and a_2. The value exactly in between is thus $(a_1 + a_2)/2$.

Question 24

Concerning an m-vector of observables $\underline{y}$ two simple hypotheses are tested against each other using the SLR-test. The hypotheses are specified as: $H_0 : \underline{y} \sim f_y(y|x = x_0)$ and $H_a : \underline{y} \sim f_y(y|x = x_a)$ with $x_a > x_0$. The observables are normally distributed, mutually uncorrelated and all have equal variance σ^2 (with σ a known value). The parameter x pertains to the location of the PDF: $E(y_i) = x$, for every $i = 1, \ldots, m$.

Let's look at the probability of incorrectly accepting H_0. What happens to it if we change the number of observations? And what if we change the standard deviation?

Solution

We don't have a sounding proof this time, but just a logical derivation. Suppose we have a small standard deviation. In this case we have very accurate measurements. We therefore will probably choose the right hypothesis.

Now suppose we have a huge number of experiments. By putting all that data together, we also get a high accuracy. So we will also (most likely) choose the right hypothesis.

However, if we want to increase the probability of incorrectly accepting H_0, we want bad measurement data. So we should increase the standard deviation (making the measurement data almost meaningless). We should also decrease the number of experiments. If we only have one useless experiment, there will be a good chance we incorrectly accept H_0.