Solutions

TAMS24/TEN1 2018-10-22

1. Let $x_1, x_2, \ldots, x_n$ be a sample of size n from the Gamma distribution given in the exercise.

   (a) The likelihoodfunction $L(\theta)$ is given by

   $$L(\theta) = \prod_{i=1}^{n} \frac{\theta(\theta x_i)^{v-1} e^{-\theta x_i}}{\Gamma(v)}.$$

   The parameter space is $\Omega_\theta = (0, \infty)$. We were not given any restrictions on θ, but without this assumption it is not clear that we end up with a density function. We form the loglikelihood and take the derivative

   $$\log L(\theta) = \sum_{i=1}^{n} \big(v \log \theta + (v-1)\log x_i - \theta x_i - \log \Gamma(v)\big),$$
   $$\frac{d \log L(\theta)}{d\theta} = \frac{n}{\theta} + \frac{n(v-1)}{\theta} - \sum_{i=1}^{n} x_i.$$

   We're seeking an extremum, so we're looking for points where the derivative is zero:

   $$\frac{n + n(v-1)}{\theta} - n\bar{x} = 0 \qquad \Leftrightarrow \qquad \theta = \frac{v}{\bar{x}},$$

   where $\bar{x}$ is the mean value of the sample. The sign-change for the derivative at the point $\widehat{\theta} = v/\bar{x}$ is $+0-$, so we're dealing with a maximum. It is also clear that $\widehat{\theta} \in \Omega_\theta$ unless all samples are equal to zero, which would be a ridiculous sample.

   (b) We replace $\bar{x}$ by the stochastic quantity $\bar{X}$, where x_j are observations of X_j that are Gamma distributed. We obtain that

   $$E\left(\frac{1}{\widehat{\Theta}}\right) = E\left(\frac{\bar{X}}{v}\right) = \frac{1}{nv}\sum_{i=1}^{n} E(X_i) = \frac{1}{nv}\frac{nv}{\theta} = \frac{1}{\theta},$$

   where we used the hint that told us that the expectation of a Gamma distributed random variable is v/θ.

   Answer: a) The ML-estimate is given by $\widehat{\theta} = v/\bar{x}$. b) See above.

2. Let H_0 be the hypothesis that the data is from a Po(1) variable X and H_1 that this is not true. In total, we have $n = 100$ observations. Suppose that H_0 is true. Then

   $$P(X = j) = \frac{1^j}{j!}e^{-1} = \frac{e^{-1}}{j!},$$

   so we can calculate the last two lines in the following table

$X = ?$	0	1	2	3	4	≥ 5
x_j (frequency)	38	33	26	2	1	0
p_j	0.368	0.368	0.184	0.061	0.015	0.004
np_j	36.8	36.8	18.4	6.1	1.5	0.4

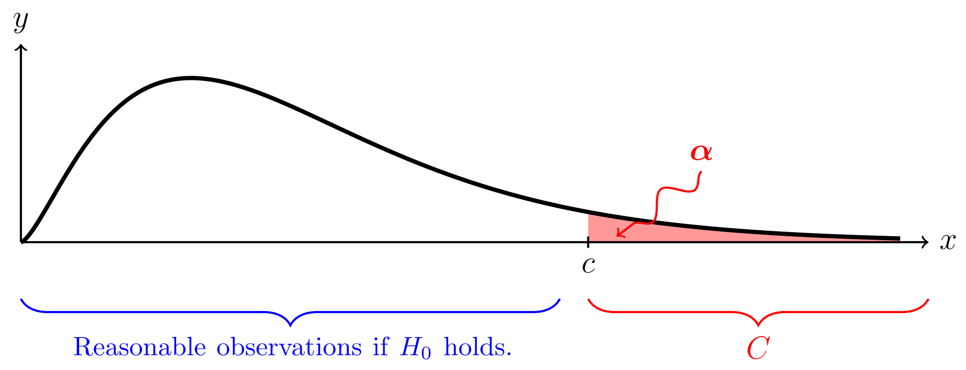
We used that $P(X \geq 5) = 1 - \sum_{j=0}^{4} p_j$ to calculate p_5. We realize here that we have a problem with the last categories since $np_j < 5$. We have to merge these to use a χ^2-test, so let us consider the following categories.

$X = ?$	0	1	2	≥ 3
x_j (frequency)	38	33	26	3
p_j	0.368	0.368	0.184	0.0805
np_j	36.8	36.8	18.4	8.05

The normal test quantity is found in

$$q = \sum_{j=0}^{3} \frac{(x_j - np_j)^2}{np_j} = \frac{(38-36.8)^2}{36.8} + \frac{(33-36.8)^2}{36.8} + \frac{(26-18.4)^2}{18.4} + \frac{(3-8.05)^2}{8.05} = 6.7387.$$

If H_0 is true, then q is an observation of $Q \overset{\text{appr.}}{\sim} \chi^2(4-1) = \chi^2(3)$. We reject H_0 if q is large, so we need a critical region C. From a table we find that $c = \chi^2_{0.01}(3) = 11.34$ and we define $C = [c, \infty)$. If $q \geq c$, we reject H_0.



Since $q \notin C$, the conclusion is that we can't reject H_0. So Sture is wrong in rejecting H_0 at this level.

Answer: Sture is wrong. The hypothesis can not be rejected at this level.

3. So the model is that for using citric acid we assume that $X_i \sim N(\mu_1, \sigma^2)$ and for acetic acid that $Y_i \sim N(\mu_2, \sigma^2)$ (same variance). All variables are assumed to be independent.

   (a) We weight together the variances according to the pooled variance:

$$s^2 = \frac{9s_1^2 + 9s_2^2}{18} = \frac{1}{2}\left(s_1^2 + s_2^2\right).$$

   It now follows that (by Cochran's and Gosset's theorems)

$$T = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{S\sqrt{\frac{1}{10} + \frac{1}{10}}} \sim t(18),$$

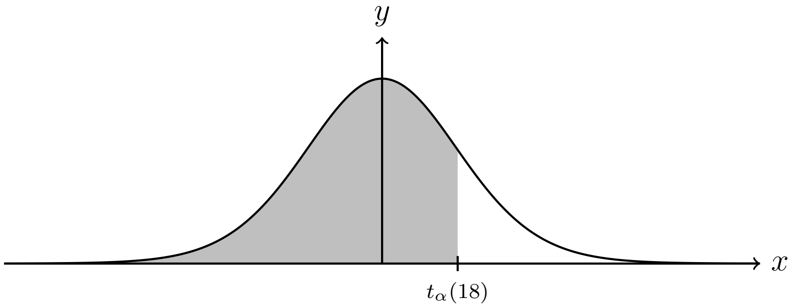
   and

$$P(T < t_\alpha(18)) = 1 - \alpha,$$

where we can solve the inequality for

$$\overline{X} - \overline{Y} - t_\alpha(18) \cdot \frac{S}{\sqrt{5}} < \mu_1 - \mu_2.$$

We use a one-sided interval since we only want to investigate if $\mu_1 > \mu_2$. From a table, we find that $t_{0.10}(18) = 1.3304$.



As an observation of S, we use $\sqrt{s^2}$, so

$$t_{0.10}(18)\frac{s}{\sqrt{5}} = 1.3304 \cdot 3.8176 = 5.0790.$$

Since $\overline{x} - \overline{y} = 8.9$, the interval is given by

$$\begin{aligned} I_{\mu_1-\mu_2} &= (8.9 - 5.0790, \infty) \\ &= (3.821, \infty). \end{aligned}$$

We see that 0 is not included in the interval, so we can claim that the citric acid produces a better yield at this significance level. (it is reasonable that $\mu_1 > \mu_2$).

(b) Let

$$H_0 : \sigma_1^2 = \sigma_2^2 = \sigma^2$$

versus

$$H_1 : \sigma_1^2 \neq \sigma_2^2.$$

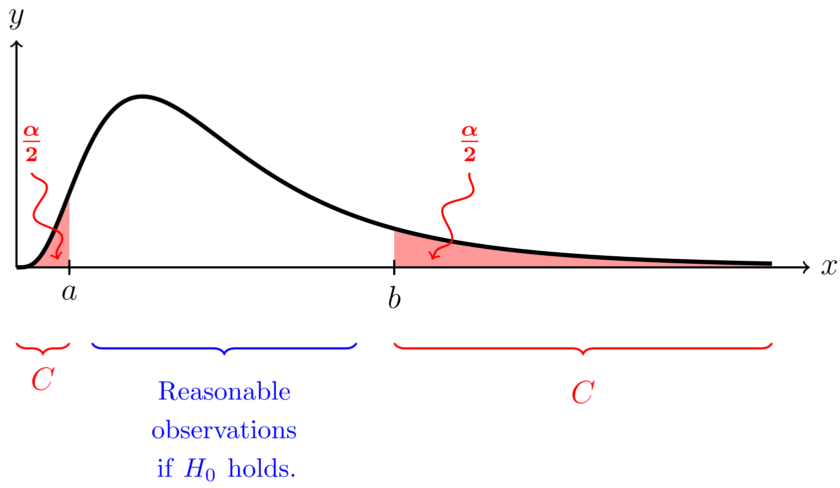
If H_0 is true, then $9S_1^2/\sigma^2 \sim \chi^2(9)$ and $9S_2^2/\sigma^2 \sim \chi^2(9)$. Since these quantities are independent, we have

$$V = \frac{\frac{9S_1^2}{\sigma^2}/(9)}{\frac{9S_2^2}{\sigma^2}/9} = \frac{S_1^2}{S_2^2} \sim F(9,9).$$

We're looking for a critical region C such that

$$\alpha = P(V \in C \mid H_0)$$

and we reject H_0 if we obtain an observation in C.



We find a and b from a table such that

$$P(V < a) = P(V > b) = \frac{\alpha}{2} = 0.025,$$

so $b = 4.0260$ and $a = 0.2484$. Since

$$v = \frac{8.5088^2}{8.5641^2} = 0.9871 \notin C$$

we can't reject H_0. The variances might be the same (it is not unreasonable).

Answer: (a) Citric acid produces a better yield. (b) Inconclusive. It is not unreasonable that the variances are the same.

4. We can formulate the problem as a matrix equation $Y = X\beta + \epsilon$, where X is the design matrix

$$X^T = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ 1 & 4 & 9 & 16 & 25 & 36 & 49 & 64 & 81 & 100 & 121 & 144 & 169 & 196 \end{pmatrix}.$$

The LSE $\widehat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ can be obtained from the well-known formula

$$(X^T X)^{-1} X^T y = \begin{pmatrix} -0.2198 \\ 0.5220 \\ 23.2692 \end{pmatrix} \cdot 10^{-3}.$$

We recognize this from the data given in the exercise (and thus it's not a step necessary for the solution). The estimated regression line can be written

$$\widehat{\mu}(x) = (-0.21978 + 0.521978x + 23.269x^2) \cdot 10^{-3}$$

and the estimate value at the point $x = 20$ is obtained as $\widehat{\mu}(20) = \widehat{\boldsymbol{\beta}}^T \cdot (1,\ 20,\ 20^2)^T \approx 9.32$, so that's the answer for the last part of this exercise. But we'll get back to that.

(a) To test if $\beta_1 = 0$, let $H_0 : \beta_1 = 0$ and $H_1 : \beta_1 \neq 0$. Assume that H_0 holds. Then

$$T = \frac{\widehat{\beta}_1 - 0}{S\sqrt{h_{11}}} \sim t(11),$$

where the distribution is clear since H_0 holds. We need a critical region C such that $P(T \in C \mid H_0) = 0.05$ and since H_1 is double sided, we choose symmetrically.

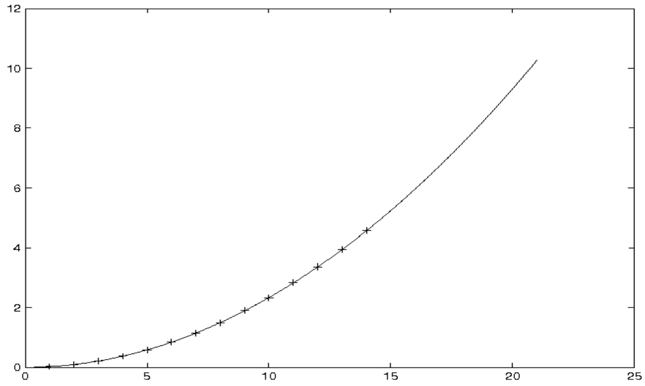
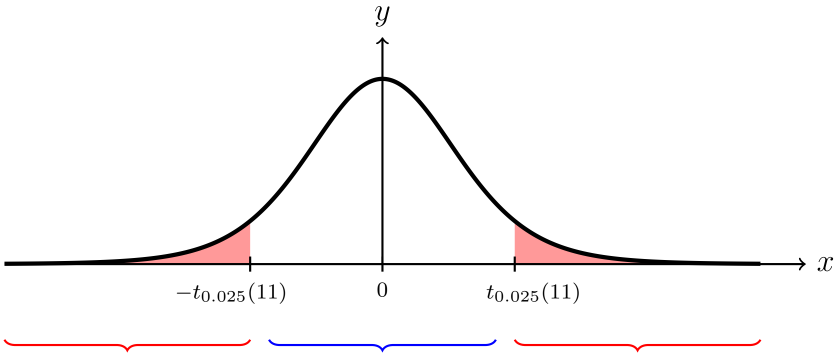


Figure 1: The given measurements match the regression line. If the model holds when $x \gg 14$ is not clear.



We find $t_{\alpha/2}(11) = t_{0.005}(11) = 3.1058$ in a table. An observation of $S\sqrt{h_{11}}$ is given by the standard error $d(\widehat{\beta}_1)$ and thus we find that the observation

$$t = \frac{0.5220 \cdot 10^{-3}}{2.0676 \cdot 10^{-3}} = 0.2525$$

does *not* belong to the critical region. So we can not reject H_0. The coefficient β_1 might very well be zero.

(b) When it comes to finding a prediction interval at $x = 11.5$, we let $u = (1\ 11.5\ 11.5^2)^T$ and Y_0 be an independent random observation at $x = 11.5$. Let $\widehat{\mu}_0$ be the estimate for μ at $x = 11.5$. A well known test quantity is

$$T = \frac{Y_0 - \widehat{\mu}_0}{S\sqrt{1 + \boldsymbol{u}^T(X^TX)^{-1}\boldsymbol{u}}} \sim t(11).$$

We can box in this variable and solve for Y_0:

$$\begin{aligned} -t < T < t \quad &\Leftrightarrow \quad -t < \frac{Y_0 - \widehat{\mu}_0}{S\sqrt{1 + \boldsymbol{u}^T(X^TX)^{-1}\boldsymbol{u}}} < t \\ &\Leftrightarrow \quad \widehat{\mu}_0 - tS\sqrt{1 + \boldsymbol{u}^T(X^TX)^{-1}\boldsymbol{u}} < Y_0 < \widehat{\mu}_0 + tS\sqrt{1 + \boldsymbol{u}^T(X^TX)^{-1}\boldsymbol{u}}, \end{aligned}$$

where $t = t_{\alpha/2}(11) = t_{0.005}(11) = 3.1058$. We can now calculate that

$$\boldsymbol{u}^T(X^TX)^{-1}\boldsymbol{u} = 0.1417,$$

so $\sqrt{1+\boldsymbol{u}^T(X^TX)^{-1}\boldsymbol{u}} = 1.0685$. As an observation of S, we use

$$s = \sqrt{\frac{5.7582 \cdot 10^{-4}}{11}} = 0.0072.$$

For $\widehat{\mu_0}$, we use the observation $\boldsymbol{u}^T\widehat{\boldsymbol{\beta}} = 3.0831$. Thus we obtain the prediction interval

$$I_{Y_0} = \Big(3.0831 \mp 3.1058 \cdot 0.0072 \cdot 1.0685\Big) = (3.0592,\ 3.1070).$$

Answer:

(a) No, β_1 could be zero. A possible interpretation is that the first degree term is drowned out in the second algorithmically, meaning that operations are always performed on all elements squarely.

(b) $(3.059,\ 3.107)$.

(c) The estimated value is 9.32. It is always uncertain to predict values outside the domain from which we've measured. In this case, there would be a gigantic shift in time consumption if the physical RAM memory would run out and the computer moves on to storing things on a hard drive instead. It might stile scale quadratically, but the constants will change. Moreover, if the hard drive runs out of space? What happens then... etc. Don't use a method outside the interval for which we have observations.

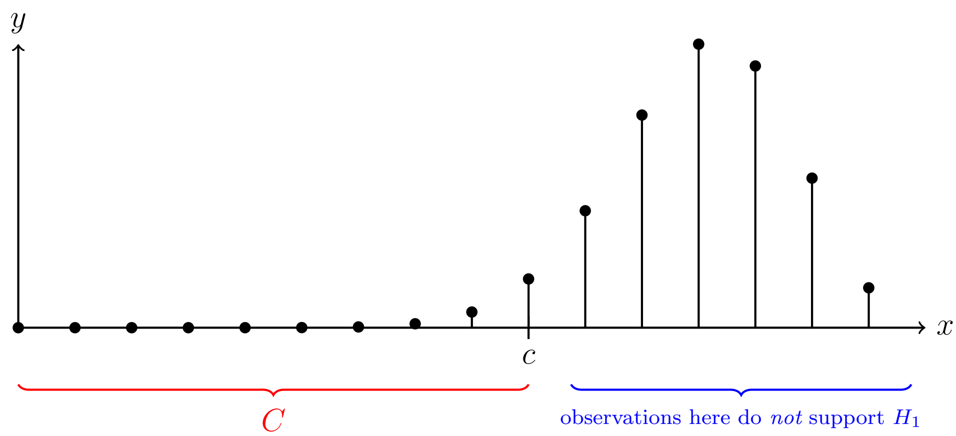
5. (a) We help Conny by performing his hypothesis test. Let X be the number of seeds that grow when planting 15. Then $X \sim \text{Bin}(n, p)$, where $n = 15$ och p is the unknown probability of a seed to grow. We want to test

$$H_0 : p = 0.8$$

versus

$$H_1 : p < 0.8.$$

Given that H_0 is true, we expect the frequency $15 \cdot 0.8 = 12$ for the number of seeds that grow. Is $x = 10$ significantly less? We need the critical region C.



Since

$$p(x) = \binom{15}{x} 0.8^x \cdot 0.2^{15-x},$$

we can calculate that

$$\sum_{x=0}^{8} p(x) = 0.0181 \quad \text{och} \quad \sum_{x=0}^{9} p(x) = 0.0611$$

so it is clear that $c = 8$ is necessary for obtaining a significance level not higher than 5% . Thus,

$$C = \{x \in \mathbf{Z} : 0 \leq x \leq 8\}$$

and our observation $x = 10 \notin C$. Hence we can't reject H_0. The seller might be speaking the truth.

(b) The power at $p = 0.65$ can be calculated straight from the definition:

$$\begin{aligned} h(0.65) &= P(H_0 \text{ rejected} \,|\, p = 0.65) = P(X \in C \,|\, p = 0.65) \\ &= \sum_{x=0}^{8} \binom{15}{x} 0.65^x \cdot 0.35^{15-x} = 0.2452. \end{aligned}$$

(c) We assume that we can approximate X by a normal distribution so that

$$X \overset{\text{appr.}}{\sim} \mathrm{N}(np, np(1-p)).$$

This will be okay if $np(1-p) \geq 10$. We'll need to check that this holds when we're done. Let $c = \Phi^{-1}(0.05)$ and $d = \Phi^{-1}(0.90)$. Then

$$0.05 = \Phi(c) = P(H_0 \text{ rejected} \,|\, p = 0.8) = P(X \leq a \,|\, p = 0.8) \approx \Phi\left(\frac{a - 0.8n}{\sqrt{n \cdot 0.8 \cdot 0.2}}\right)$$

$$\Leftrightarrow \quad c = \frac{a - 0.8n}{\sqrt{n \cdot 0.8 \cdot 0.2}} \quad \Leftrightarrow \quad a = c\sqrt{n \cdot 0.16} + 0.8n$$

and

$$0.9 = P(H_0 \text{ rejected} \,|\, p = 0.65) = P(X \leq a \,|\, p = 0.65) \approx \Phi\left(\frac{a - 0.65n}{\sqrt{n \cdot 0.65 \cdot 0.35}}\right)$$

$$\Leftrightarrow \quad d = \frac{a - 0.65n}{\sqrt{n \cdot 0.2275}} \quad \Leftrightarrow \quad a = d\sqrt{n \cdot 0.2275} + 0.65n$$

Thus

$$c\sqrt{n \cdot 0.16} + 0.8n = d\sqrt{n \cdot 0.2275} + 0.65n \quad \Leftrightarrow \quad \sqrt{n} = \frac{d\sqrt{0.2275} - c\sqrt{0.16}}{0.15} = 8.4614,$$

so $n = 72$ is sufficient. With this n we can find a according to

$$a = d\sqrt{n \cdot 0.2275} + 0.65n = 51.99$$

or

$$a = c\sqrt{n \cdot 0.16} + 0.8n = 52.02.$$

We choose $a = 51$ to be sure that the critical region doesn't become too large. Since $n = 72$ makes $np(1-p) > 10$ for both $p = 0.8$ and $p = 0.65$, our approximation should be okay.

Doing an exact verification, we can see that if $X \sim Bin(72, p)$ we obtain that

$$P(X \leq 51 \,|\, p = 0.8) = 0.0406 \quad \textit{and} \quad P(X \leq 51 \,|\, p = 0.65) = 0.8783.$$

Alternate interpretation. One could also interprete the question as using the exactly same significance level we ended up with earlier, i.e., $\alpha = 0.0181$. In this case, letting $c = \Phi^{-1}(0.0181) = -2.0947$ we will obtain that $\sqrt{n} = 9.6609$, so $n = 94$ would be chosen. This leads to $a = 67$.

Answer:

(a) We can't reject H_0.

(b) The power is 0.2452.

(c) $n = 72$ (giving $c = 51$).

6. It is clear that $D = \text{cov}(\boldsymbol{Y})$ is a diagonal matrix since the components are independent. Moreover, since $A^2 = A^T = A$, we can see that

$$\boldsymbol{Y}^T A \boldsymbol{Y} = \boldsymbol{Y}^T A^T A \boldsymbol{Y} = (A\boldsymbol{Y})^T (A\boldsymbol{Y}),$$

and similarly that $\boldsymbol{Y}^T B \boldsymbol{Y} = (B\boldsymbol{Y})^T (B\boldsymbol{Y})$. Let us show that $A\boldsymbol{Y}$ and $B\boldsymbol{Y}$ are independent. The result will then follow since $\boldsymbol{Y}^T A \boldsymbol{Y}$ and $\boldsymbol{Y}^T B \boldsymbol{Y}$ clearly are functions of $A\boldsymbol{Y}$ and $B\boldsymbol{Y}$, respectively. Since $\boldsymbol{Y}$ is normally distributed, this is also true for $A\boldsymbol{Y}$ and $B\boldsymbol{Y}$. Thus it is enough to show that the variables are uncorrelated to obtain independence. The covariance is given by

$$\begin{aligned}
\text{cov}(A\boldsymbol{Y}, B\boldsymbol{Y}) &= E(A\boldsymbol{Y}(B\boldsymbol{Y})^T) - E(A\boldsymbol{Y})E(B\boldsymbol{Y})^T \\
&= AE(\boldsymbol{Y}\boldsymbol{Y}^T)B^T - AE(\boldsymbol{Y})(BE(\boldsymbol{Y}))^T \\
&= A\left(E(\boldsymbol{Y}\boldsymbol{Y}^T) - E(\boldsymbol{Y})E(\boldsymbol{Y})^T\right)B^T = A\left(\text{cov}(\boldsymbol{Y})\right)B^T = ADB = DAB = 0,
\end{aligned}$$

if $D = \sigma^2 I$.

Answer: See above.