MATH11400 Statistics 1 2010–11

Homepage http://www.stats.bris.ac.uk/%7Emapjg/Teach/Stats1/

Solution Sheet 10

1. **Model assumptions**: (a) The lifetimes of the 10 tyres are a simple random sample from the population of lifetimes for all tyres currently produced by that company. (b) The population distribution for those lifetimes is $N(\mu, \sigma^2)$, where both μ and σ are unknown.

   Hypotheses: H_0: $\mu = 42$ versus H_1: $\mu < 42$.
   The null hypothesis H_0 corresponds to *no difference* between the actual mean of the population of lifetimes for that company's tyres and the claimed mean lifetime of $42(\times 1000)$ miles. In practice, this claim would probably be interpreted as a claim that the mean lifetime was *at least* 42, for which the alternative hypothesis H_1 would corresponds to a mean lifetime less than 42.

   Test Statistic: Since $\bar{X}$ is the natural estimator of μ, we base our test statistic on $\bar{X} - \mu_0 = \bar{X} - 42$. Since σ^2 is unknown, we take as our test statistic $T(X_1, \ldots, X_n) = \sqrt{n}(\bar{X} - \mu_0)/S$ which has a t_{n-1} distribution when H_0 is true (i.e. when $\mu = \mu_0 = 42$).

   For the given data, $n = 10$, $\bar{x} = 41$, and $s^2 = 12.89$, so the observed test statistic is $t_{obs} = \sqrt{10}(41 - 42)/\sqrt{12.89} = -0.8808$. Also, since $n = 10$, $T \sim t_9$ when H_0 is true.,

   ***p*-value**: Since the alternative of interest is H_1: $\mu < 42$, the values of T which are less consistent with H_0 than t_{obs} are the set of values $\{T < t_{obs}\}$ so

   p-value $=$ $P(T < t_{obs} | H_0 \text{ true}) = P(t_9 < -0.8808)$.
   Using **R** or the hint in the question, $P(t_9 < -0.8808) =$
   `1-pt(0.8808,9)=1-0.79933=0.20067.`

   Critical region: Since the alternative of interest is H_1: $\mu < 42$, the values of T which are less consistent with H_0 than t are the set of values $\{T < t\}$. Thus the critical region of values for which the test would reject H_0 is of the form $C = \{T < c^*\}$. A test has significance level α if P(Reject $H_0 | H_0$ true) $= \alpha$. Thus, for a 0.05-level test, c^* is defined by the condition

   $$0.05 = \alpha = P(\text{Reject } H_0 | H_0 \text{ true}) = P(T < c^* \mid H_0 \text{ true}) = P(t_9 < c^*)$$

   so by symmetry $P(t_9 > -c^*) = 0.05$ and$-c^* = t_{9;0.05} = 1.833$ and the resulting critical region is $C = \{T < -1.833\}$ which does not contain t_{obs}.

   Conclusions: A p-value of 0.2 is relatively large, so there is little evidence that H_0 is not true. The observed test statistic value $t_{obs} = -0.8808$ falls well outside the critical region of the 0.05-level test, so we would accept H_0, and conclude that there is no reason to reject the manufacturer's claim that the mean lifetime of their tyres is equal to $42,000$ miles.

2. (a)
   Model assumptions: (a) The 52 measured blood sugar levels are a simple random sample from the population of blood sugar levels for all pregnant women in their third trimester of pregnancy. (b) The population distribution is $N(\mu, 10^2)$, where μ is unknown.

Notes (i) In practice, one might want to look carefully at how these women were chosen for observation, to be sure that they really were a representative simple random sample from the population of pregnant women and that there was no common systematic factor that might have affected their blood glucose level. (ii) Even if the distribution of blood glucose level in the population of pregnant women does not actually follow a Normal distribution, the central limit theorem and the sample size of 52 means the distribution of $\bar{X}$ will be close to a Normal distribution, especialy since the true distribution is likely to be unimodal. (iii) It is often realistic to assume to assume that the difference we are investigating will have had more effect (if any) on the population mean than on the population variability, in which case it may be reasonable to assume the variance for the population under consideration is the same as that for the reference population.

Hypotheses: H_0: $\mu = 80$ versus H_1: $\mu < 80$.
The null hypothesis H_0 corresponds to *no difference* between the mean blood sugar level for all pregnant women in their third trimester of pregnancy and the known mean blood sugar level for healthy women who are not pregnant. The alternative hypothesis H_1 corresponds to the mean for pregnant women being lower than that for healthy women who are not pregnant.

Test Statistic: Since $\bar{X}$ is the natural estimator of μ, we base our test statistic on $\bar{X} - \mu_0 = \bar{X} - 80$. Since the population standard deviation $\sigma_0 = 10$ is known and $n = 52$, we can take as our test statistic $T(X_1, \ldots, X_n) = \sqrt{n}(\bar{X} - \mu_0)/\sigma_0 = \sqrt{52}(\bar{X} - 80)/10$, where $\bar{X} \sim N(\mu, \sigma_0^2/n) = N(\mu, 10^2/25)$.
Thus, when H_0 is true (i.e. when $\mu = \mu_0 = 80$) we have $T = \sqrt{52}(\bar{X} - 80)/10 \sim N(0, 1)$.

The data give $\bar{x} = 70.12$ so the observed test statistic is $t_{obs} = -7.1246$.

***p*-value**: Since the alternative of interest is H_1: $\mu < 80$, the values of T which are less consistent with H_0 than t_{obs} are the set of values $\{T < t_{obs}\}$ so

$$p\text{-value} = P(T < t_{obs}|H_0 \text{ true}) = P(Z < -7.1246) \;\; [\text{where } \; Z \sim N(0,1)] = \Phi(-7.1246).$$

R gives `5.204034e-13`, so the p-value is tiny!

Conclusions: The p-value is so small that there is overwhelming evidence that the data are not consistent with H_0 being true. We would reject H_0 in favour of H_1, and conclude that the mean blood sugar level for the population of pregnant women in their third trimester of pregnancy is lower than that for the population of healthy womenwho are not pregnant.

(b)
Critical Region for $\boldsymbol{\alpha = 0.01}$: Now assume the test was carried out using a procedure with significance level $\alpha = 0.01$. Since the alternative of interest is H_1: $\mu < 80$, the values of T which are less consistent with H_0 than t are the set of values $\{T < t\}$. Thus the critical region of values for which the test would reject H_0 is of the form $C = \{T < c^*\}$. A test has significance level α if P(Reject $H_0|H_0$ true) $= \alpha$. Thus, for a 0.01-level test, c^* is defined by the condition

$$\begin{aligned} 0.01 &= \alpha = P(\text{Reject } H_0|H_0 \text{ true}) = P(T < c^* \mid H_0 \text{ true}) \\ &= P(Z < c^*) \;\; [\text{where } \; Z \sim N(0,1)] = \Phi(c^*), \\ \text{so } c^* &= \Phi^{-1}(0.01) = \texttt{qnorm(0.01)} = -2.326 \end{aligned}$$

and the resulting critical region is $C = \{T \leq -2.326\}$.

Type II error: Assume now we are interested in the specific alternative H_1: $\mu = 79$. When H_1 is true $\bar{X} \sim N(79, 10^2/52)$, so $\sqrt{52}(\bar{X} - 79)/10 \sim N(0, 1)$.

Now P(Type II error) = P(Accept $H_0 | H_1$ true) = P($T > c^* \,|\, H_1$ true) = P($T > -2.326 \,|\, H_1$ true). But $T > -2.326 \Leftrightarrow \sqrt{52}(\bar{X} - 80)/10 > -2.326 \Leftrightarrow \bar{X} > 76.77 \Leftrightarrow \sqrt{52}(\bar{X} - 79)/10 > -1.605$. Thus

$$\begin{aligned} P(\text{Type II error}) &= P(T > -2.326 \,|\, H_1 \text{ true}) = P(\sqrt{52}(\bar{X} - 79)/10 > -1.605 |\, H_1 \text{ true}) \\ &= P(Z > -1.605) = P(Z < 1.605) = \texttt{pnorm(1.605)} = 0.946 \end{aligned}$$

Finally, the power of the test to discriminate between H_0 and H_1 is $1-$ P(Type II error) = 0.054. This power is not very large – even if H_1 was true the test procedure would only detect the fact that it was true in 1 in 20 samples of the sort taken here.

3. The `t.test()` command has default significance level $\alpha = 0.05$. Thus the test procedure has a probability 0.05 of rejecting H_0 when in fact H_0 is true.

   Here, you are taking samples from the Normal distribution with your specified mean μ_0 and then testing whether the sample looks like it comes from the distribution with mean μ_0. Even though μ_0 is the correct mean for the distribution from which the sample was drawn, the 0.05-level test will reject the null hypothesis H_0: $\mu = \mu_0$ in $100\alpha\% = 5\% = 1$ in 20 of your samples. Of course the actual number of rejections in each batch of twenty samples you take will vary due to the random nature of the sampling process, but you should see about one (false) rejection for every 20 samples you take.

   You can alter the procedure so that it gives fewer (false) rejections when H_0 is true by decreasing the significance level – but only at the expense of increasing the type II error for value of μ under the alternative hypothesis.

4. **Model assumptions**: We assume that, for given $x_1, \ldots, x_n$, the Y_i are independent Normally distributed random variables with mean $\alpha + \beta x_i$ and variance σ^2.

   Least squares estimates: Summary statistics for the data set are:
   $n = 5 \quad \sum x_i = 21 \quad \sum y_i = 12 \quad \sum x_i^2 = 111 \quad \sum y_i^2 = 46 \quad \sum y_i x_i = 69.$
   From these we get $n = 5$ (so $n - 2 = 3$), $\bar{x} = 4.2$, $\bar{y} = 2.4$, $ss_{xx} = \sum x_i^2 - (\sum x_i)^2/n = 22.8$, $ss_{yy} = \sum y_i^2 - (\sum y_i)^2/n = 17.2$ and $ss_{xy} = \sum x_i y_i - (\sum x_i \sum y_i)/n = 18.6$. Thus the least squares estimates for α and β are

   $$\hat{\beta} = ss_{xy}/ss_{xx} = 0.8158 \qquad \hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = -1.0263$$

   while the estimates of σ^2 and of $\text{Var}(\hat{\alpha})$ and $\text{Var}(\hat{\beta})$ are $\hat{\sigma}^2 = \dfrac{(ss_{yy} - ss_{xy}^2/ss_{xx})}{(n-2)} = 0.6754$, $s_{\hat{\alpha}}^2 = \hat{\sigma}^2(1/n + \bar{x}^2/ss_{xx}) = 0.6577 = 0.8110^2$, $s_{\hat{\beta}}^2 = \hat{\sigma}^2/ss_{xx} = 0.02962 = 0.1721^2$.

   (a) **Hypothesis test for** α: Here the hypotheses to be tested are H_0: $\alpha = 0$ versus H_1: $\alpha \neq 0$. Note that the null hypothesis corresponds to the fitted regression line passing through the origin – in most cases this will not be either a plausible or an interesting hypothesis.

   Test Statistic: As in question 1, under these model assumptions $(\hat{\alpha} - \alpha)/s_{\hat{\alpha}}$ has the t-distribution with $n - 2$ degrees of freedom. Thus we take as test statistic $T = \hat{\alpha}/s_{\hat{\alpha}}$

which has the t_{n-2} (i.e. the t_3) distribution when H_0 is true (since then $\alpha = 0$). This gives $t_{obs} = \hat{\alpha}/s_{\hat{\alpha}} = -1.0263/0.8110 = -1.265$.

***p*-value**: The alternative of interest is H_1: $\alpha \neq 0$, so the values of T which are less consistent with H_0 than t_{obs} are the set $\{|T| > |t_{obs}|\}$. Thus the p-value = $P(|T| > |t_{obs}||H_0 \text{ true}) = P(|t_3| > 1.265) = 2(1 - P(t_3 < 1.265))$. The command `pt(1.265,3)` in **R** gives $P(t_3 < 1.265) = 0.8524$, giving a p-value of 0.2952.

Critical region: Since the values of T which are less consistent with H_0 than a value t are the set $\{|T| > |t|\}$, the critical region of values for which an γ-level test would reject H_0 is of the form $C = \{|T| > c^*\}$, where c^* is defined by the condition: $\gamma = P(\text{Reject } H_0|H_0 \text{ true}) = P(|T| > c^* \,|\, H_0 \text{ true}) = 2P(t_3 > c^*)$, i.e. $P(t_3 > c^*) = \gamma/2$. Thus, for $\gamma = 0.05$, $c^* = t_{3;0.025} =$ `qt(0.975,3)` $= 3.182$ giving $C = \{|T| > 3.182\}$. Note that here C does not contain t_{obs}.

Conclusions: The p-value is not that small and the observed test statistic $t_{obs} = -1.265$ is not in the critical region of the 0.05-level test, so there is no evidence that we should reject H_0 in favour of H_1. We therefore conclude that there is no evidence to reject the hypothesis H_0: $\alpha = 0$ that the fitted regression line does pass through the origin.

(b) **Hypothesis test for** β: Here the hypotheses of interest are H_0: $\beta = 0$ versus H_1: $\beta \neq 0$. Now the null hypothesis H_0 corresponds to the fitted regression line being parallel to the x-axis, i.e. to the hypothesis that the mean value of Y does not vary with the value of the predictor variable x.

Test Statistic: Under the model assumptions $(\hat{\beta} - \beta)/s_{\hat{\beta}}$ has the t-distribution with $n - 2$ degrees of freedom. Thus we take as test statistic $T = \hat{\beta}/s_{\hat{\beta}}$ which has the t_{n-2} distribution when H_0 is true (since then $\beta = 0$). This gives $t_{obs} = \hat{\beta}/s_{\hat{\beta}} = 0.8158/0.1721 = 4.740$.

***p*-value**: The alternative of interest is H_1: $\beta \neq 0$, so again the values T less consistent with H_0 than t_{obs} are the set $\{|T| > |t_{obs}|\}$. Thus the p-value = $P(|T| > |t_{obs}||H_0 \text{ true}) = P(|t_3| > 4.740) = 2(1 - P(t_3 < 4.740))$. If you use the command `pt(4.740,3)` in **R**, it gives $P(t_3 < 4.740) = 0.99110$.

Critical region: Exactly the same argument as for α above gives that the critical region of values for which an γ-level test would reject H_0 is of the form $C = \{|T| > c^*\}$, where $P(t_3 > c^*) = \gamma/2$. Thus, for $\gamma = 0.05$, $c^* = t_{3;0.025} = 3.182$ giving $C = \{|T| > 3.182\}$. Note that this time C does contain t_{obs}.

Conclusions: The p-value is very small and the observed test statistic value $t_{obs} = 4.740$ falls well within the critical region of the 0.05-level test, so there is strong evidence that we should reject H_0: $\beta = 0$ in favour of H_1: $\beta \neq 0$ and conclude that the mean value of Y does vary with the value of the predictor x.

5. **Model assumptions**: Let Y_i denote the log metabolic rate and let x_i denote the log body mass for the ith dog. From the question we can assume that, for given $x_1, \ldots, x_n$, the Y_i are independent Normally distributed random variables with mean $\alpha + \beta x_i$ and variance σ^2.

   Least squares estimates: From the summary statistics we get $n = 7$ (so $n - 2 = 5$), $\bar{x} = 2.543$, $\bar{y} = 6.513$, $ss_{xx} = \sum x_i^2 - (\sum x_i)^2/n = 3.975$,

$ss_{yy} = \sum y_i^2 - (\sum y_i)^2/n = 1.5056$ and $ss_{xy} = \sum x_i y_i - (\sum x_i \sum y_i)/n = 2.4339$.
Thus the least squares estimates for α and β are

$$\hat{\beta} = ss_{xy}/ss_{xx} = 0.6124 \qquad \hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 4.9558$$

while the estimates of σ^2 and of $\mathrm{Var}(\hat{\alpha})$ and $\mathrm{Var}(\hat{\beta})$ are

$$\hat{\sigma}^2 = \frac{(ss_{yy} - ss_{xy}^2/ss_{xx})}{(n-2)} = 0.003033,$$

$$s_{\hat{\alpha}}^2 = \hat{\sigma}^2(1/n + \bar{x}^2/ss_{xx}) = 0.005367, s_{\hat{\beta}}^2 = \hat{\sigma}^2/ss_{xx} = 0.0007630.$$

Confidence interval for α: From §9.6 of your notes, under these model assumptions $(\hat{\alpha} - \alpha)/s_{\hat{\alpha}}$ has the t-distribution with $n-2$ degrees of freedom, so from §9.2 of your notes the end points (c_L, c_U) of a $100(1-\gamma)\%$ confidence interval for α are given by

$$c_L = \hat{\alpha} - t_{n-2;\gamma/2} \times s_{\hat{\alpha}} \qquad \text{and} \qquad c_U = \hat{\alpha} + t_{n-2;\gamma/2} \times s_{\hat{\alpha}}.$$

For a 99% confidence interval, $100(1-\gamma) = 99$ so $\gamma/2 = 0.005$ and from **R**, $t_{5;0.005} =$ `qt(0.995,5)=4.032`. Thus the 99% confidence interval for α has end points

$$c_L = 4.9558 - 4.032 \times 0.07326 = 4.661 \qquad c_U = 4.9558 + 4.032 \times 0.07326 = 5.251.$$

Confidence interval for β: In the same way, under these model assumptions $(\hat{\beta}-\beta)/s_{\hat{\beta}}$ has the t-distribution with $n-2$ degrees of freedom, so the end points (c_L, c_U) of a $100(1-\gamma)\%$ confidence interval for β are given by

$$c_L = \hat{\beta} - t_{n-2;\gamma/2} \times s_{\hat{\beta}} \qquad \text{and} \qquad c_U = \hat{\beta} + t_{n-2;\gamma/2} \times s_{\hat{\beta}}.$$

For a 99% confidence interval, $100(1-\gamma) = 99$ so $\gamma/2 = 0.005$ and from **R**, $t_{5;0.005} =$ `qt(0.995,5)=4.032`. Thus the 99% confidence interval for β has end points

$$c_L = 0.6124 - 4.032 \times 0.02762 = 0.501 \qquad c_U = 0.6124 + 4.032 \times 0.02762 = 0.724.$$

6. **Model assumptions**: Let Y_i denote the autumn rainfall and let x_i denote the observed spring rainfall for the ith year. From the question, we can assume that, for given $x_1, \ldots, x_n$, the Y_i are independent Normally distributed random variables with mean $\alpha + \beta x_i$ and variance σ^2.

   Least squares estimates: The summary statistics for the data set are:
   $n = 10$ $\sum x_i = 49.4$ $\sum y_i = 66.9$ $\sum x_i^2 = 311.22$ $\sum y_i^2 = 540.49$ $\sum y_i x_i = 400.26$.

   From the summary statistics we get: $\bar{x} = 4.94$, $\bar{y} = 6.69$, $ss_{xx} = \sum x_i^2 - (\sum x_i)^2/n = 67.184$, $ss_{yy} = \sum y_i^2 - (\sum y_i)^2/n = 92.929$ and $ss_{xy} = \sum x_i y_i - (\sum x_i \sum y_i)/n = 69.774$.

   Thus the least squares estimate for β is $\hat{\beta} = ss_{xy}/ss_{xx} = 1.039$ while the estimates of σ^2 and $\mathrm{Var}(\hat{\beta})$ are $\hat{\sigma}^2 = (ss_{yy} - ss_{xy}^2/ss_{xx})/(n-2) = 2.558$ and $s_{\hat{\beta}}^2 = \hat{\sigma}^2/ss_{xx} = 0.03808$.

(a) **Confidence interval for** β: As in question 1, under these model assumptions $(\hat{\beta} - \beta)/s_{\hat{\beta}}$ has the t-distribution with $n-2$ degrees of freedom (here t_8), so the end points (c_L, c_U) of a $100(1-\gamma)\%$ confidence interval for β are given by $c_L = \hat{\beta} - t_{n-2;\gamma/2} \times s_{\hat{\beta}}$ and $c_U = \hat{\beta} + t_{n-2;\gamma/2} \times s_{\hat{\beta}}$. For a 90% confidence interval, $100(1-\gamma) = 90$ so $\gamma/2 = 0.05$ and from **R**, $t_{8;0.05}$ =`qt(0.95,8)=1.860`. Thus the 90% confidence interval has end points $c_L = 1.039 - 1.86 \times 0.1951 = 0.674$, $c_U = 1.039 + 1.86 \times 0.1951 = 1.404$.

(b) **Hypothesis test for** β: Here the hypotheses of interest are H_0: $\beta = 0$ versus H_1: $\beta \neq 0$. As in question 2, H_0 corresponds to the hypothesis that the mean value of Y does not vary with the value of the predictor variable x.

Test Statistic: Under the model assumptions $(\hat{\beta} - \beta)/s_{\hat{\beta}}$ has the t-distribution with $n-2$ degrees of freedom. Thus we take as test statistic $T = \hat{\beta}/s_{\hat{\beta}}$ which has the t_{n-2} distribution when H_0 is true (since then $\beta = 0$). This gives $t_{obs} = \hat{\beta}/s_{\hat{\beta}} = 1.039/0.1951 = 5.324$.

***p*-value**: The alternative of interest is H_1: $\beta \neq 0$, so again the values of T which are less consistent with H_0 than t_{obs} are the set $\{|T| > |t_{obs}|\}$. Thus the p-value = $P(|T| > |t_{obs}||H_0 \text{ true}) = P(|t_8| > 5.324) = 2(1 - P(t_8 < 5.324))$. **R** gives `pt(5.324,8)` $= 0.9996462$ from which we obtain the p-value as $2(1 - 0.9996462) = 0.0007$.

Critical region: Exactly the same argument as for question 2 above gives that the critical region of values for which an test with significance level 0.10 would reject H_0 is of the form $C = \{|T| > c^*\}$, where $c^* = t_{8;0.05} = 1.860$ giving $C = \{|T| > 1.860\}$. Here $t_{obs} \in C$.

Conclusions: The p-value is extremely small and the observed test statistic value $t_{obs} = 5.324$ falls well within the critical region of the 0.10-level test, so there is very strong evidence that we should reject H_0 in favour of H_1 and conclude that the mean autumn rainfall does vary with the value of the rainfall for the previous spring.

(c) The γ-level test of H_0: $\beta = 0$ against H_1: $\beta \neq 0$ accepts $H_0 \iff |t_{obs}| \leq t_{n-2;\gamma/2}$ $\iff -t_{n-2;\gamma/2} \leq \hat{\beta}/s_{\hat{\beta}} \leq t_{n-2;\gamma/2} \iff \hat{\beta} - t_{n-2;\gamma/2}s_{\hat{\beta}} \leq 0$ and $\hat{\beta} + t_{n-2;\gamma/2}s_{\hat{\beta}} \geq 0 \iff$ $c_L \leq 0 \leq c_U \iff$ the $100(1-\gamma)\%$ confidence interval for β contains the value $\beta = 0$.

7. The summary statistics for this data set are:
$n = 7$ $\sum x_i = 44$ $\sum y_i = 9.6$ $\sum x_i^2 = 344$ $\sum y_i^2 = 13.36$ $\sum y_i x_i = 57$.

From the summary statistics we get: $\bar{x} = 6.2857$, $\bar{y} = 1.3714$, $ss_{xx} = \sum x_i^2 - (\sum x_i)^2/n = 67.429$, $ss_{yy} = \sum y_i^2 - (\sum y_i)^2/n = 0.19429$ and $ss_{xy} = \sum x_i y_i - (\sum x_i \sum y_i)/n = -3.3429$.

Thus the least squares estimates for α and β are $\hat{\beta} = ss_{xy}/ss_{xx} = -0.04958$ and $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} = 1.683$, while the estimates of σ^2, $\text{Var}(\hat{\alpha})$ and $\text{Var}(\hat{\beta})$ are $\hat{\sigma}^2 = (ss_{yy} - ss_{xy}^2/ss_{xx})/(n-2) = 0.005712$, $s_{\hat{\alpha}}^2 = \hat{\sigma}^2(1/n + \bar{x}^2/ss_{xx}) = 0.004163$ and $s_{\hat{\beta}}^2 = \hat{\sigma}^2/ss_{xx} = 0.00008471$.

We assume the data are observations satisfying the simple linear Normal regression model $Y_i = \alpha + \beta x_i + e_i$, $i = 1, \ldots, n$, where the e_i are i.i.d. $N(0, \sigma^2)$.

(a) The values on the line beginning `(Intercept)` are: (i) $\hat{\alpha}$ (the estimate of α),
(ii) $s_{\hat{\alpha}}$ (the standard error, which estimates the standard deviation $\hat{\alpha}$),

(iii) $t_{obs} = \hat{\alpha}/s_{\hat{\alpha}}$ (the observed test statistic for testing H_0: $\alpha = 0$ vs. H_1: $\alpha \neq 0$),
(iv) $P(|W| \geq |t_{obs})$, where $W \sim t_{n-2}$ (the p-value of the data for the test).

The values on the line beginning `pig.littersize` are the corresponding quantities for estimating or testing hypotheses about β (i.e. (i) $\hat{\beta}$, (ii) $s_{\hat{\beta}}$, (iii) $t_{obs} = \hat{\beta}/s_{\hat{\beta}}$, and (iv) $P(|W| \geq |t_{obs})$, where $W \sim t_{n-2}$.

The values on the line beginning `Residual` are $\hat{\sigma}$ and $n - 2$.

The line beginning `Signif. codes` tells us that three asterisks $\star\star\star$ indicates that the corresponding p-value lies between 0.001 and 0, while two asterisks $\star\star$ indicates that the corresponding p-value lies between 0.01 and 0.001.

(b) The output indicates that, for a test of H_0: $\beta = 0$ vs. H_1: $\beta \neq 0$, the p-value is 0.00297. Since this is extremely small, we would reject H_0 and conclude that the mean weight of the pigs in a given litter does vary with the number of pigs in the litter.

8. The sea-level and high-altitude times for each runner are likely to be dependent – those who did well at sea level also did well at high altitude. However, the difference in the race time for each runner may well be independent of the difference for other runners, and these differences may well have the same variability. Thus we use a **paired t-test**.

   Model assumptions: For $i = 1, \ldots, 8$ let X_i denote the sea-level time for the ith runner, let Y_i denote the high-altitude time for the same runner, and let $W_i = X_i - Y_i$. We assume $W_1, \ldots, W_8$ are a simple random sample from the $N(\delta, \sigma^2)$ distribution, where δ and σ^2 are unknown.

   Hypotheses: H_0: $\delta = 0$ versus H_A: $\delta < 0$.
   The null hypothesis H_0 corresponds to a mean of zero for the differences between the race times at sea level and high altitude. The alternative hypothesis of interest is that race times at high altitude are systematically larger than those at sea level, i.e. that this mean is negative.

   Test Statistic: As in §9.12 of your notes, we take as our test statistic $T = \sqrt{n}\bar{W}/\hat{\sigma}_W$. Here $n = 8$, $\hat{\sigma}_W^2 = S_W^2 = \sum_{i=1}^{8}(W_i - \bar{W})^2/(8-1)$, and T has the t_7 distribution when H_0 is true.

   For the given data, the w_i values are $-2.1, 0.3, -1.6, -2.0, 1.1, -3.4, 0.1, -1.8$, with mean $\bar{w} = -1.175$ and variance $s_W^2 = 2.290714$. so the observed test statistic is $t_{obs} = -2.1958$.

   ***p*-value**: The alternative of interest is H_A: $\delta < 0$, so values of T which are less consistent with H_0 than t_{obs} are the set $\{T < t_{obs}\}$. Thus the p-value = $P(T < t_{obs}|H_0 \text{ true})$ = $P(t_7 < -2.1958)$ = $P(t_7 > 2.1958)$ (by symmetry) = $1 - P(t_7 < 2.1958)$. If you use the command `pt(2.1958,7)` in **R**, it gives $P(t_7 < 2.1958) = 0.96794$ and a p-value of 0.03206.

   Critical region: Since the values of T which are less consistent with H_0 than a value t are the set $\{T < t\}$, the critical region of values for which an α-level test would reject H_0 is of the form $C = \{T < c^*\}$, where c^* is defined by the condition: $\alpha = P(\text{Reject } H_0 | H_0 \text{ true}) = P(T < c^* \,|\, H_0 \text{ true}) = P(t_7 < c^*) = P(t_7 > -c^*)$. Thus, for $\alpha = 0.05$, $-c^* = t_{7;0.05} =$ `qt(0.95,7)` $= 1.895$ and $c^* = -1.895$ giving $C = \{T < -1.895\}$.

Conclusions: The p-value is very small, so there is strong evidence that H_0 is not true. The observed test statistic value $t_{obs} = -2.1958$ falls well within the critical region of the 0.05-level test. Thus we would reject H_0 in favour of H_A, and conclude that the mean race times at high altitude are indeed larger than those at sea level.

9. Here we have two entirely independent samples of salaries, and we are told in the question that we may assume that the population variances are the same in the private and public sectors, so it is appropriate to use a **pooled two-sample t-test** .

   Model assumptions: Let $X_1, \ldots, X_9$ denote the salararies for the nine private sector posts and let $Y_1, \ldots, Y_{10}$ denote the salararies for the ten public sector posts. We assume: $X_1, \ldots, X_9$ are a simple random sample from the $N(\mu_X, \sigma_X^2)$ distribution; $Y_1, \ldots, Y_{10}$ are a simple random sample from the $N(\mu_Y, \sigma_Y^2)$ distribution; $\sigma_X^2 = \sigma_Y^2 = \sigma^2$ (say); and the two samples are independent of each other.

   Hypotheses: H_0: $\mu_X - \mu_Y = 0$ versus H_A: $\mu_X - \mu_Y > 0$.
   The null hypothesis corresponds to there being no systematic difference between the privare sector and the public sector salaries; the alternative hypothesis corresponds to private sector salaries being systematically higher, i.e. to a positive difference between the means.

   Test Statistic: As in §9.11 and §9.11.1 of your notes, for a pooled t-test we take as test statistic $T = (\bar{X} - \bar{Y})/S_p\sqrt{1/n + 1/m}$, where $S_p^2 = [(n-1)S_X^2 + (m-1)S_Y^2]/(n+m-2)$, and T has the t_{n+m-2} distribution (i.e. $T \sim t_{17}$) when H_0 is true.

   Here $n = 9$, $\sum x_i = 104.0$, $\sum x_i^2 = 1218.92$, $\bar{x} = 11.5556$, $s_x^2 = 2.1428$,
   $m = 10$, $\sum y_i = 102.3$, $\sum y_i^2 = 1071.69$, $\bar{y} = 10.2300$, $s_y^2 = 2.7957$,
   so $s_p^2 = 2.4884$ and the observed test statistic is $t_{obs} = 1.8289$.

   ***p*-value**: The alternative of interest is H_A: $\mu_X - \mu_Y > 0$, so the values of T which are less consistent with H_0 than t_{obs} are the set $\{T > t_{obs}\}$. Thus the p-value = $P(T > t_{obs}|H_0 \text{ true})$ = $P(t_{17} > 1.8289)$ = $1 - P(t_{17} < 1.8289)$. If you use the command `pt(1.8289,17)` in **R**, it gives $P(t_{17} < 1.8289) = 0.9575$ and a p-value of 0.0425.

   Critical region: The values of T less consistent with H_0 than a value t are the set $\{T > t\}$, so the critical region of values for which an α-level test would reject H_0 is of the form $C = \{T > c^*\}$, where c^* is defined by the condition: $\alpha = P(\text{Reject } H_0 | H_0 \text{ true}) = P(T > c^* \,|\, H_0 \text{ true}) = P(t_{17} > c^*)$. For $\alpha = 0.05$, $c^* = t_{17;0.05} =$ `qt(0.95,17)` $= 1.740$ giving $C = \{T > 1.740\}$.

   Conclusions: The p-value is quite small and gives reasonable evidence that H_0 is not true. The observed test statistic value $t_{obs} = 1.8289$ falls just within the critical region of the 0.05-level test. Thus we would reject H_0 in favour of H_A, and conclude that private sector starting salaries are significantly higher than public sector starting salaries.

10. As in Question 8, we will use a **paired t-test**.

    Model assumptions: For $i = 1, \ldots, 10$ let W_i denote the difference between these 'before' and 'after' usage for the ith household, and assume $W_1, \ldots, W_{10}$ are a simple random sample from the $N(\delta, \sigma^2)$ distribution, where δ and σ^2 are unknown.

    Hypotheses: H_0: $\delta = 0$ versus H_A: $\delta > 0$,
    since the alternative hypothesis of interest is that usage has been reduced.

Test Statistic: $T = \sqrt{n}\bar{W}/\hat{\sigma}_W$. Here $n = 10$, $\hat{\sigma}_W^2 = S_W^2 = \sum_{i=1}^{10}(W_i - \bar{W})^2/(10-1)$ and T has the t_9 distribution when H_0 is true.

For the given data, $\bar{w} = 56$ and $s_W^2 = 10937.78$, so $t_{obs} = 1.6933$.

***p*-value**: For H_A: $\delta > 0$, the values of T which are less consistent with H_0 than t_{obs} are $\{T > t_{obs}\}$. Thus the p-value = $P(T > t_{obs}|H_0$ true) = $P(t_9 > 1.6933)$ = $1 - P(t_9 < 1.6933)$. If you use the command `pt(1.6933,9)` in **R**, it gives a $P(t_9 < 1.6933) = 0.93768$ and a p-value of 0.06232.

Critical region: Here, the critical region is of the form $C = \{T > c^*\}$, where, for a 0.05-level test, c^* is defined by the condition: $0.05 = P(T > c^* \mid H_0 \text{ true}) = P(t_9 > c^*)$, so $c^* = t_{9;0.05} = 1.833$, giving $C = \{T > 1.833\}$.

Conclusions: The p-value is not very small and the observed test statistic value $t_{obs} = 1.6933$ does not fall within the critical region of the 0.05-level test. Thus there is not enough evidence to reject H_0 in favour of H_A, and hence not enough evidence to conclude that the monitor is effective at reducing electrical consumption.

11. We will use **Welch's test**, since we have two independent samples but σ_X^2 and σ_Y^2 may differ.

Model assumptions: Let $X_1, \ldots, X_{10}$ denote the relative changes for the ten subjects with the calcium supplement and let $Y_1, \ldots, Y_{11}$ denote the relative changes for the eleven subjects with the placebo. We assume: $X_1, \ldots, X_{10}$ are a simple random sample from the $N(\mu_X, \sigma_X^2)$ distribution; $Y_1, \ldots, Y_{11}$ are a simple random sample from the $N(\mu_Y, \sigma_Y^2)$ distribution; and the two samples are independent of each other.

Hypotheses: H_0: $\mu_X - \mu_Y = 0$ versus H_A: $\mu_X - \mu_Y < 0$.
Here, reduction in blood pressure corresponds to negative relative changes. The alternative hypothesis corresponds the reduction for subjects with the calcium supplement being, on average, greater than that for subjects using the placebo.

Test Statistic: From §9.11.2 of your notes, Welch's test uses $T = (\bar{X} - \bar{Y})/\sqrt{S_X^2/n + S_Y^2/m}$, where T has approximately the t_ν distribution when H_0 is true,
and where $\nu = (S_X^2/n + S_Y^2/m)^2/((S_X^2/n)^2/(n-1) + (S_Y^2/m)^2/(m-1))$.

Here, $n = 10,\ m = 11,\ \bar{x} = -5,\ s_x^2 = 76.444,\ \bar{y} = 0.2727,\ s_y^2 = 34.818$,
so $\nu = 15.591$ and the observed test statistic is $t_{obs} = -1.6037$.

***p*-value**: The alternative of interest is H_A: $\mu_X - \mu_Y < 0$, p-value = $P(T < t_{obs}|H_0$ true) = $P(t_{15.591} < -1.6037)$. As ν is non-integer, you can use `pt(-1.6037,15.591)` in **R**, which gives the p-value as $P(t_{15.591} < -1.6037) = 0.06442$.

Conclusions: The p-value is similar to that of Question 5. Again, there is not really enough evidence to reject H_0 in favour of H_A. Thus we would conclude that there is not enough evidence to reject the hypothesis that the calcium supplement has no more effect than using a placebo.