A short quiz

1. A new staff to the school was told that there is only $\frac{1}{10{,}000}$ chance that the lift in the Maths building will be faulty, causing the passenger(s) to be trapped. Suppose that the staff goes to work 5 days a week for 42 weeks per year and use the lift twice a day. Assume that the outcome (of lift failure) each day is independent.

   (a) What is the probability that the staff will *not* be trapped during the first year?

   (b) What is the expected number of times the staff will be trapped in 5 years?

2. A linear model was fitted to the Tooth Growth data using R. The code and output is given below. The response is the length of odontoblasts (cells responsible for tooth growth) in 60 guinea pigs. Each animal received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods, orange juice (OJ) or ascorbic acid (a form of vitamin C and coded as VC).

```r
library(tidyverse)
data("ToothGrowth")
lm1 <- lm(len ~ dose*supp, data=ToothGrowth)
summary(lm1)
```

```
##
## Call:
## lm(formula = len ~ dose * supp, data = ToothGrowth)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.2264 -2.8462  0.0504  2.2893  7.9386
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   11.550      1.581   7.304 1.09e-09 ***
## dose           7.811      1.195   6.534 2.03e-08 ***
## suppVC        -8.255      2.236  -3.691 0.000507 ***
## dose:suppVC    3.904      1.691   2.309 0.024631 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.083 on 56 degrees of freedom
## Multiple R-squared:  0.7296, Adjusted R-squared:  0.7151
## F-statistic: 50.36 on 3 and 56 DF,  p-value: 6.521e-16
```

   (a) Write down the fitted model for each of the supplement type.

   (b) Calculate the 95% confidence interval for the **difference** in the mean tooth length for the VC and OJ supplement types, at the does level of 1.17.

3. Consider the matrix form of the linear model $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$. Derive the least square estimator of $\boldsymbol{\beta}$.

Answers

1. (a) Let X denote the number of times the staff is trapped during the first year. We have $X \sim B(480, 0.0001)$. Then $P(X=0) = \binom{480}{0} 0.0001^0 0.9999^{480} = 0.9531$.

   (b) In this case, we have $X \sim B(2400, 0.0001)$. Hence $E(X) = 2400 \times \frac{1}{10{,}000} = 0.24$ times.

2. (a) For OJ type, $\texttt{len} = 11.55 + 7.811 * \texttt{dose}$.
   For VC type, $\texttt{len} = 3.295 + 11.715 * \texttt{dose}$.

   (b) The CI for $\mu_{\mathrm{VC}} - \mu_{\mathrm{OJ}}$ is $(-5.81, -1.59)$.

3. The least square estimator of $\boldsymbol{\beta}$ is obtained by minimising the sum of squared errors (SSE) with respect to $\boldsymbol{\beta}$:

$$
\begin{aligned}
\mathrm{SSE} &= \boldsymbol{\epsilon}^\top \boldsymbol{\epsilon} = (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})^\top (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}) \\
\frac{\partial \mathrm{SSE}}{\partial \boldsymbol{\beta}} &= -2\,\boldsymbol{X}^\top (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}) = 0 \\
(\boldsymbol{X}\boldsymbol{X}^\top)^{-1}\boldsymbol{\beta} &= \boldsymbol{X}^\top \boldsymbol{Y} \\
\Rightarrow \quad \hat{\boldsymbol{\beta}} &= (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{Y}
\end{aligned}
$$

provided $\boldsymbol{X}^\top \boldsymbol{X}$ has an inverse.