2.1 Problem 1

(a) This was done in class. Let $T \subseteq \{1, \ldots, n\}$ be the set of indices corresponding to the true null (wlog, we assume that $T$ is non-null, otherwise no rejection can be false). Let $A$ be the event that there is at least one false rejection and $B$ the event $\mathcal{H}_T$ is rejected. If $i \in T$ is rejected, then that means by the closure principle procedure that $\mathcal{H}_T$ is also rejected (because $\{i\} \subseteq T$). Hence $A \subseteq B$. We conclude that

$$\mathbb{P}(A) \leq \mathbb{P}(B) \leq \alpha,$$

where we used that the test for $\mathcal{H}_T$ is a test at level $\alpha$.

(b) A simple test for model selection in regression analysis is the $F$-test. Consider $I \subseteq \{1, \ldots, p\}$, we denote $\text{RSS}_I$, the residual sum of squares of the reduced model with the constraints $\beta_i = 0, i \in I$, and $\text{RSS}_{\text{full}}$ is the residual sum of squares of the full model (including all $p$ variables). Then the $F$-statistic is given by

$$F_I = \frac{(\text{RSS}_I - \text{RSS}_{\text{full}})/|I|}{\text{RSS}_{\text{full}}/(n-p)}.$$

Under the null hypothesis $\mathcal{H}_I$ (i.e., $\beta_i = 0, i \in I$) and with the extra assumption that the data is normally distributed with common variance, then the $F$-statistic follows the $F$-distribution $F_{|I|, n-p}$ and we can define an $\alpha$-level test $\{\varphi_I = 1\} = \{F_I \geq F^1_{|I|, n-p}\}$ (without the normal distribution, we could evaluate the threshold through bootstrap etc.).

Another suggestion from the students: using Bonferroni, i.e., $\{\varphi_I = 1\} = \{\min_{i \in I} p_i \leq \alpha/|I|\}$, where $p_i$ is the p-value of for example the t-test of $\mathcal{H}_i$ (i.e., $\beta_i = 0$). However this test is much more conservative than a $F$-tests.

(c) Denote $R_0 \subseteq R$ the subset of null variables in $R$ (hence $\tau(R) = |R_0|$), and $T \subseteq \{1, \ldots, p\}$ the subset of all null variables. First notice that if $\tau(R) > t_\alpha(R)$, then by definition of $t_\alpha(R)$, we have $R_0 \in \mathcal{X}$: hence $\mathcal{H}_{R_0}$ is rejected and by the closing principle procedure ($R_0 \subseteq T$), we must have $\mathcal{H}_T$ rejected, which happens with probability less or equal to $\alpha$. In summary,

$$\mathbb{P}(\tau(R) \leq t_\alpha(R)) = 1 - \mathbb{P}(\tau(R) > t_\alpha(R)) \geq 1 - \mathbb{P}(T \in \mathcal{X}) \geq 1 - \alpha.$$
(d) The stepwise procedure select a subset $R \subseteq \{1, \ldots, p\}$. Using the local tests described in part (b) as well as the closure principle, we can compute $t_\alpha(R)$ as defined in part (c) which an upper confidence limit correct with probability at least $1 - \alpha$.

(e) Now we are given $R' \subset R$. We can similarly take $t_\alpha(R')$.

(f) Consider $T \subseteq \{1, \ldots, p\}$ the set of all null variables and $E$ the event that $\mathcal{H}_T$ is rejected. On $E^c$, $T$ is not rejected and by the closure principle, any $R \subseteq \{1, \ldots, p\}$, we have $T \cap R \subseteq R$ and therefore $T \cap R$ is not rejected. By definition, we must then have $t_\alpha(R) = \max\{\lvert I \rvert : I \subseteq R, I \not\in \mathcal{X} \} \geq \lvert R \cap T \rvert = \tau(R)$. Hence,

$$\mathbb{P}(\tau(R) \leq t_\alpha(R), \forall R) \geq \mathbb{P}(E^c) \geq 1 - \alpha.$$  

Hence, we deduce that all the upper confidence limits are valid simultaneously with probably at least $1 - \alpha$.

(g) One of the advantages of this method is that it is very general: it does not depend on the way the set $R$ is selected. Furthermore, the upper confidence bounds hold simultaneously over every subset $R$. An obvious limit is computation: in order to compute the outcome of the closure principle, one need to perform $2^p$ tests. However if the local test uses Bonferroni, this procedure can be run more efficiently (see lectures). Lastly, since the method does not rely on the actual subset $R$ that ends up being selected, it may overestimate the actual number of false rejections $\tau(R)$, by lack of adaptivity.

### 2.2 Problem 2

For this question, I used the article Antoine et. al. “Early uneven ear input induces long-lasting differences in left–right motor function”. In: *PLoS Biology* 16.3 (March 2018). Retrieved from: https://doi.org/10.1371/journal.pbio.2002988

In the article, the authors perform nineteen experiments and test a hypothesis separately for each experiment. For each experiment, they reject the corresponding hypothesis exactly when the $p$ value is less than 0.05, and do not address the multiple-testing problem.

In the following, I randomly choose 10 of the reported p-values and perform Hochberg’s test on these 10 $p$-values.

```r
## [1] 0.0019 0.0001 0.6700 0.0450 0.0460 0.9400 0.0380 0.0039
    0.0007 0.4200
```

The $p$-values that end up rejected by Hochberg’s procedure are:
They correspond to the following values:

We see that Hochberg’s test rejects four hypotheses. In the paper, the authors reject these hypotheses as well as four additional hypotheses with \( p \)-values below 0.05. In particular, their procedure does not control the family-wise error rate at the 0.05 level.

### Problem 3

(a) We have that, for any \( i \in H_0 \),

\[
\left( \frac{V_i}{1 \lor R} \right)^2 = \frac{V_i}{1 \lor R^2} = \sum_{k=1}^{n} \frac{V_i 1\{R=k\}}{k^2} = \sum_{k=1}^{n} \frac{1(p_i \leq qk/n) 1\{R(p_i \to 0)=k\}}{k^2}.
\]

Taking expectation conditional on \( p \setminus p_i \) thus yields that, since \( 1(p_i \leq qk/n) \) is independent of \( p \setminus p_i \) while \( 1\{R(p_i \to 0)=k\} \) is \( p \setminus p_i \)-measurable,

\[
\mathbb{E} \left[ \left( \frac{V_i}{1 \lor R} \right)^2 \big| p \setminus p_i \right] = \sum_{k=1}^{n} \mathbb{E}(1(p_i \leq qk/n) 1\{R(p_i \to 0)=k\} | p \setminus p_i) \frac{1}{k^2} = \sum_{k=1}^{n} \frac{\Pr(p_i \leq qk/n) 1\{R(p_i \to 0)=k\}}{k^2} = \sum_{k=1}^{n} \frac{q}{n} \frac{1\{R(p_i \to 0)=k\}}{k}.
\]

Taking expectations of both sides hence yields

\[
\mathbb{E} \left( \frac{V_i}{1 \lor R} \right)^2 = \frac{q}{n} \sum_{k=1}^{n} \frac{\Pr(R(p_i \to 0) = k)}{k} = q\pi_0 \frac{1}{n_0} \mathbb{E} \left( \frac{1}{R(p_i \to 0)} \right).
\]

2-3
For $i \neq j \in \mathcal{H}_0$, we have that
\[
\left( \frac{V_i}{1 \lor R} \right) \left( \frac{V_j}{1 \lor R} \right) = \frac{VV_j}{1 \lor R^2} = \sum_{k=2}^{n} \frac{V_i V_j 1_{\{R=k\}}}{k^2} = \sum_{k=2}^{n} \frac{1_{\{p_i \leq qk/n\}} 1_{\{p_j \leq qk/n\}} 1_{\{R(p_i \rightarrow 0, p_j \rightarrow 0) = k\}}}{k^2}.
\]
Taking expectation conditional on $p \setminus (p_i, p_j)$ thus yields that, since $1_{\{p_i \leq qk/n\}}$ and $1_{\{p_j \leq qk/n\}}$ are independent of $p \setminus (p_i, p_j)$ and of each other, while $1_{\{R(p_i \rightarrow 0, p_j \rightarrow 0) = k\}}$ is $p \setminus (p_i, p_j)$-measurable,
\[
\mathbb{E}\left[\left( \frac{V_i}{1 \lor R} \right) \left( \frac{V_j}{1 \lor R} \right) \mid p \setminus (p_i, p_j)\right] = \sum_{k=2}^{n} \frac{\Pr(p_i \leq qk/n) \Pr(p_j \leq qk/n) 1_{\{R(p_i \rightarrow 0, p_j \rightarrow 0) = k\}}}{k^2} = \sum_{k=2}^{n} \frac{q^2}{n^2} 1_{\{R(p_i \rightarrow 0, p_j \rightarrow 0) = k\}} = \frac{q^2}{n^2}.
\]
Hence, we have that
\[
\mathbb{E}\left( \frac{V_i}{1 \lor R} \right) \left( \frac{V_j}{1 \lor R} \right) = \frac{q^2}{n^2}.
\]
Now, we can compute
\[
\text{Var}(\text{FDP}) = \mathbb{E}\text{FDP}^2 - (\mathbb{E}\text{FDP})^2
= \mathbb{E}\left( \sum_{i \in \mathcal{H}_0} \frac{V_i}{1 \lor R} \right)^2 - \frac{\pi_0^2 q^2}{n^2}
= \sum_{i \in \mathcal{H}_0} \mathbb{E}\left( \frac{V_i}{1 \lor R} \right)^2 + 2 \sum_{i < j} \mathbb{E}\left( \frac{V_i}{1 \lor R} \right) \left( \frac{V_j}{1 \lor R} \right) - \frac{\pi_0^2 q^2}{n^2}
= q \pi_0 \frac{1}{n_0} \sum_{i \in \mathcal{H}_0} \mathbb{E}\left( \frac{1}{R(p_i \rightarrow 0)} \right) + n_0 (n_0 - 1) \frac{q^2}{n^2} - \frac{\pi_0^2 q^2}{n^2}
= q \pi_0 \text{Ave}_{i \in \mathcal{H}_0} \mathbb{E}\left( \frac{1}{R(p_i \rightarrow 0)} \right) - \frac{n_0 q^2}{n^2}
\leq q \pi_0 \text{Ave}_{i \in \mathcal{H}_0} \mathbb{E}\left( \frac{1}{R(p_i \rightarrow 0)} \right).
\]
We see from the above analysis that if $n_0 = 0$ or $n \to \infty$, then the inequality is an equality and so the bound is tight.
(b) Since \( n_0 q^2 / n^2 \) is very small in any reasonable situation, the approximation in part (a) is very good.

Now, for \( i \in H_0 \), we expect that \( H_i \) is not rejected, but if that we were to set \( p_i \) to 0, it would not change the BH threshold significantly. Hence, we would expect that we would roughly reject the \( R \) initial hypotheses as well as now rejecting \( H_i \). That is, \( R(p_i \to 0) \approx R + 1 \).

If we believe that there are only a few non-nulls, then it is reasonable to approximate \( \pi_0 \approx 1 \). Hence, we might approximate the variance by

\[
\hat{\text{Var}}(\text{FDP}) = \frac{q}{R + 1}.
\]

(c) By taking a reasonable range of the realised FDP to be the mean plus or minus two standard deviations, we would expect the FDP to fall in

\[
\left( q - 1.96 \sqrt{\frac{q}{R + 1}}, q + 1.96 \sqrt{\frac{q}{R + 1}} \right) \approx (0.01902, 0.18098).
\]

(d) As already discussed, for small number of rejections, we can estimate \( \pi_0 \) with 1 as small number of rejections indicate that there are very few non-nulls or equivalently very large number of nulls.

For large number of rejections, if we have all the p-values with us, we use the following estimate for \( n_0 \) as discussed in class, \( \hat{n}_0 = 2(n - R(.5)) \) which in turn gives the estimate for \( \pi_0 \).

For reasons already given above, we obtain the estimate for \( E_{R(p_i \to 0)} \frac{1}{R+1} \) as \( \frac{1}{R+1} \) for all \( i \in H_0 \) which in turn gives the estimate for \( \text{Var}(\text{FDP}) \).

A reasonable estimate for FDR would be \( \pi_0 q \). So a realistic range(95\% confidence interval) for realized FDP would have two sided limits as

\[
(\max \{0, \text{FDR} - 1.96 \times \text{Var}(\text{FDP})\}, \text{FDR} + 1.96 \times SD(\text{FDP}))
\]

(Here we have assumed that \( \frac{\text{FDP} - \text{FDR}}{\sqrt{\text{Var}(\text{FDP})}} \) is standard Gaussian)

One sided limit would be \( (0, \text{FDR} + 1.644 \times SD(\text{FDP})) \).

We run simulations for two different scenarios and use the above strategy to obtain a reasonable range for FDP depending on the situation. We use \( q=0.1 \) in both cases.

**CASE I** We simulate the \( n=500 \) dimensional multivariate Normal with mean vector chosen as follows: First \( k=60 \) elements of the vector are \( (1.05 \times \sqrt{2\log(n)}) \) and rest are 0 and the Covariance matrix is chosen to be the Identity matrix. So, we have 440 many true nulls, and mostly we have small number of rejections, so we assume \( \pi_0 \approx 1 \) and we find that both one sided and two sided limits cover FDP for approximately 96\% times when simulations were repeated \( N=500 \) times. [In this case, mean number of rejections over 500 simulations turned out to be 51, sample mean of FDP over 500 simulations is 0.09 while mean of the FDR estimate is 0.1, sample variance estimate of FDP is 0.00216 while mean of the variance estimate used is 0.001923] Below is the
histogram plot for FDP over 500 simulations which somewhat validates the Normality assumption.

CASE II We simulate the n=500 dimensional multivariate Normal with mean vector chosen as follows: First k=250 elements of the vector are \((1.05 \times \sqrt{2 \log(n)})\) and rest are 0 and the Covariance matrix is chosen to be the Identity matrix. So, we have 250 many true nulls, and we have large number of rejections, so we assume 
\[
\pi_0 \approx \frac{2(n-\bar{R}(0.5))}{n}
\]
and we find that both one sided and two sided limits cover FDP for approximately 94% times when simulations were repeated N=500 times. [In this case, mean number of rejections over 500 simulations turned out to be 229, sample mean of FDP over 500 simulations is 0.0503 while mean of the FDR estimate is 0.0514, sample variance estimate of FDP is 0.000218 while mean of the variance estimate used is 0.000227]

Below is the histogram plot for FDP over 500 simulations which somewhat validates the Normality assumption.

We observe that appropriately adjusting for the estimates depending on the situation is giving a good estimate for the range of FDP.