Problem 1(30 points)

(a) If there are just two categories, the test statistic can be written as:

\[
T = \frac{(X_1 - np_{01})^2}{np_{01}} + \frac{(X_2 - np_{02})^2}{np_{02}}
\]

\[
= \frac{(X_1 - np_{01})^2}{np_{01}} + \frac{(n - X_1 - n + np_{01})^2}{n(1 - p_{01})}
\]

\[
= \frac{(X_1 - np_{01})^2}{np_{01}} + \frac{(X_1 - np_{01})^2}{n(1 - p_{01})}
\]

\[
= \frac{(X_1 - np_{01})^2}{n} \left[ \frac{1}{p_{01}} + \frac{1}{1 - p_{01}} \right]
\]

\[
= \frac{(X_1 - np_{01})^2}{np_{01}(1 - p_{01})}.
\]

Now by Central Limit Theorem under the null hypothesis,

\[
\frac{\sqrt{n}(X_1/n - p_{01})}{\sqrt{p_{01}(1 - p_{01})}} \xrightarrow{d} N(0, 1).
\]

Let \( Z = \frac{\sqrt{n}(X_1/n - p_{01})}{\sqrt{p_{01}(1 - p_{01})}} \), then \( Z \xrightarrow{d} N(0, 1) \) and

\[
Z = \frac{\sqrt{n}(X_1/n - p_{01})}{\sqrt{p_{01}(1 - p_{01})}} = \frac{X_1 - np_{01}}{\sqrt{np_{01}(1 - p_{01})}}.
\]

Therefore, we notice that \( T = Z^2 \xrightarrow{d} \chi^2_1 \).

(b) Since \( \sum_{i=1}^{d} p_{0i} = 1 \), we eliminate \( p_{0d} \) and write it in terms of the other probabilities. Hence we consider \( \tilde{p} = (p_{01}, \ldots, p_{0(d-1)}) \). We know by using MLE asymptotics that, under the null hypothesis,

\[
\sqrt{n} \left( \begin{bmatrix} X_1/n \\ X_2/n \\ \vdots \\ X_{d-1}/n \end{bmatrix} - \tilde{p} \right) \xrightarrow{d} N \left( 0, \begin{bmatrix} p_{01}(1 - p_{01}) & -p_{01}p_{02} & \cdots & -p_{01}p_{0(d-1)} \\ -p_{01}p_{02} & p_{02}(1 - p_{02}) & \cdots & -p_{02}p_{0(d-1)} \\ \vdots & \vdots & \ddots & \vdots \\ -p_{01}p_{0(d-1)} & -p_{02}p_{0(d-1)} & \cdots & p_{0(d-1)}(1 - p_{0(d-1)}) \end{bmatrix} \right)
\]

Now we can show that the inverse of the covariance matrix, let us call it \( \Sigma \) is given by

\[
I(p) = \Sigma^{-1} = \begin{bmatrix}
\frac{1}{p_{01}} + \frac{1}{p_{02}} & \frac{1}{p_{02}} & \cdots & \frac{1}{p_{02}} \\
\frac{1}{p_{02}} & \frac{1}{p_{02}} + \frac{1}{p_{0d}} & \cdots & \frac{1}{p_{0d}} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{p_{0d}} & \frac{1}{p_{0d}} & \cdots & \frac{1}{p_{0(d-1)} + 1/p_{0d}}
\end{bmatrix}
\]
In order to show this, we show that

\[ \Sigma \Sigma^{-1} = I_{d-1}. \]

Now,

\[
\Sigma \Sigma^{-1} = \begin{bmatrix}
    p_{01}(1 - p_{01}) & -p_{01}p_{02} & \cdots & -p_{01}p_{0(d-1)} \\
    -p_{01}p_{02} & p_{02}(1 - p_{02}) & \cdots & -p_{02}p_{0(d-1)} \\
    \vdots & \vdots & \ddots & \vdots \\
    -p_{01}p_{0(d-1)} & -p_{02}p_{0(d-1)} & \cdots & p_{0(d-1)}(1 - p_{0(d-1)})
\end{bmatrix} \begin{bmatrix}
    \frac{1}{p_{01}} + \frac{1}{p_{0d}} & \frac{1}{p_{02}} + \frac{1}{p_{0d}} & \cdots & \frac{1}{p_{0d}} \\
    \frac{1}{p_{0d}} & \frac{1}{p_{0d}} & \cdots & \frac{1}{p_{0(d-1)}} + \frac{1}{p_{0d}}
\end{bmatrix}.
\]

So, for \( i \neq j \)

\[
(\Sigma \Sigma^{-1})_{ij} = -\sum_{k=1}^{d-1} \frac{p_{0k}p_{0k}}{p_{0d}} + \frac{p_{0i}(1 - p_{0i})}{p_{0d}} - p_{0i}p_{0j} \left( \frac{1}{p_{0j}} + \frac{1}{p_{0d}} \right)
\]

\[
= -\frac{d-1}{p_{0d}} + \frac{p_{0i}}{p_{0d}} - p_{0i} = 0,
\]

\[
(\Sigma \Sigma^{-1})_{ii} = -\sum_{k=1}^{d-1} \frac{p_{0k}p_{0k}}{p_{0d}} + \frac{p_{0i}(1 - p_{0i})}{p_{0d}} \left( \frac{1}{p_{0i}} + \frac{1}{p_{0d}} \right)
\]

\[
= -\frac{d-1}{p_{0d}} + \frac{1}{p_{0d}} + \frac{p_{0i}}{p_{0d}}
\]

\[
= -\frac{p_{0i}(1 - p_{0d})}{p_{0d}} + \frac{1}{p_{0d}} + \frac{p_{0i}}{p_{0d}} = 1.
\]

Therefore,

\[ \Sigma \Sigma^{-1} = I_{d-1}. \]

Now as,

\[
\sqrt{n} \left( \begin{bmatrix} X_1/n \\ X_2/n \\ \vdots \\ X_{d-1}/n \end{bmatrix} - \bar{\mu} \right)^d \xrightarrow{d} N(0, \Sigma)
\]

and \( I(p) = \Sigma^{-1} \),

\[
n \left( \begin{bmatrix} X_1/n \\ X_2/n \\ \vdots \\ X_{d-1}/n \end{bmatrix} - \bar{\mu} \right)^T I(p) \left( \begin{bmatrix} X_1/n \\ X_2/n \\ \vdots \\ X_{d-1}/n \end{bmatrix} - \bar{\mu} \right) \xrightarrow{d} \chi^2_{d-1}.
\]
Now, the LHS can be simplified into

\[
LHS = n \left( \begin{bmatrix} X_1/n \\ X_2/n \\ \vdots \\ X_{d-1}/n \end{bmatrix} - \bar{p} \right)^T \left( \begin{bmatrix} \frac{1}{p_01} + \frac{1}{p_0d} & \frac{1}{p_12} + \frac{1}{p_0d} & \cdots & \frac{1}{p_{d-1}d} + \frac{1}{p_0d} \\ \frac{1}{p_01} & \frac{1}{p_12} + \frac{1}{p_0d} & \cdots & \frac{1}{p_{d-1}d} + \frac{1}{p_0d} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{p_01} & \frac{1}{p_12} & \cdots & \frac{1}{p_{d-1}d} + \frac{1}{p_0d} \end{bmatrix} \right) \left( \begin{bmatrix} X_1/n \\ X_2/n \\ \vdots \\ X_{d-1}/n \end{bmatrix} - \bar{p} \right)
\]

\[
= n \sum_{i=1}^{d-1} \sum_{j=1}^{d-1} \frac{(X_i - np_{0i})}{n} \frac{(X_j - np_{0j})}{n} (\Sigma^{-1})_{ij}
\]

\[
= \sum_{i=1}^{d-1} \sum_{j=1}^{d-1} \frac{(X_i - np_{0i})(X_j - np_{0j})}{np_{0d}} + \sum_{i=1}^{d-1} \frac{(X_i - np_{0i})^2}{np_{0i}}.
\]

Now,

\[
(X_d - np_{0d})^2 = \left( n - \sum_{i=1}^{d-1} X_i - n + n \sum_{i=1}^{d-1} p_{0i} \right)^2
\]

\[
= \left( \sum_{i=1}^{d-1} (X_i - np_{0i}) \right)^2
\]

\[
= \sum_{i=1}^{d-1} \sum_{j=1}^{d-1} (X_i - np_{0i})(X_j - np_{0j}).
\]

Therefore,

\[
LHS = n \left( \begin{bmatrix} X_1/n \\ X_2/n \\ \vdots \\ X_{d-1}/n \end{bmatrix} - \bar{p} \right)^T I(p) \left( \begin{bmatrix} X_1/n \\ X_2/n \\ \vdots \\ X_{d-1}/n \end{bmatrix} - \bar{p} \right) = T \overset{d}{\sim} \chi^2_{d-1}.
\]

**Alternate proof for finding** \( I(p) \): Say if we wanted to derive the inverse matrix \( (\Sigma^{-1}) \), then we first notice that \( \Sigma \) can be written as:

\[
\Sigma = A - vv^T, \quad A = \begin{bmatrix} p_{01} & 0 & \cdots & 0 \\ 0 & p_{02} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_{0(d-1)} \end{bmatrix}, \quad v = \begin{bmatrix} p_{01} \\ p_{02} \\ \vdots \\ p_{0(d-1)} \end{bmatrix}.
\]

Then by using the Sherman–Morrison formula, we get

\[
\Sigma^{-1} = A^{-1} + A^{-1}vv^T A^{-1} \frac{1}{1 - v^T A^{-1} v}, \quad \text{where} \quad A^{-1} = \begin{bmatrix} \frac{1}{p_{01}} & 0 & \cdots & 0 \\ 0 & \frac{1}{p_{02}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{p_{0(d-1)}} \end{bmatrix}.
\]
The test statistic we consider is

We have

Problem 2 (20 points)

Now under the null hypothesis, of the standard normal distribution. Therefore for a particular statistic

Let the observed value of

Therefore,

and

Plugging everything in we get,

Problem 2 (20 points)

We have \( X_1, \ldots, X_n \sim N(\theta, 1) \) and we would like to test,

\[ H_0 : \theta = 0 \quad \text{versus} \quad H_1 : \theta > 0. \]

The test statistic we consider is \( T_n = \overline{X} \). Therefore under \( H_0, \theta = 0 \) and

\[ \sqrt{n}T_n \sim N(0, 1). \]

Let the observed value of \( T_n(x_1, \ldots, x_n) = t_n \). Then the p-value now can be given by

\[ p = P(T_n \geq t_n) = P(\sqrt{n}T_n \geq \sqrt{n}t_n) = 1 - \Phi(\sqrt{n}t_n) = \Phi(-\sqrt{n}t_n). \]

Therefore for a particular statistic \( T_n \), p-value is given by \( p = \Phi(-\sqrt{n}T_n) \), where \( \Phi \) is the CDF of the standard normal distribution.

Now under the null hypothesis,

\[
P(p \leq u) = P(\Phi(-\sqrt{n}T_n) \leq u)
= P(\sqrt{n}T_n \leq \Phi^{-1}(u)) \quad \text{(Since, } \Phi \text{ is continuous and increasing.)}
= P(\sqrt{n}T_n \geq -\Phi^{-1}(u))
= 1 - \Phi(-\Phi^{-1}(u)) = \Phi(\Phi^{-1}(u)) = u.
\]

Therefore, \( p \) is uniformly distributed.

**Note:** The proof doesn’t directly use the normality assumption, that is, the CDF of Normal could have been replaced by CDF of any distribution and the proof would still hold. The proof only depends on the CDF being continuous and increasing.
Problem 3 (20 points)

(a) The p-value is defined as the smallest $\alpha$ at which we would reject $H_0$. That is, the p-value is given by,

$$ p = \inf \{ \alpha : \{X_1, \ldots, X_n\} \in R_\alpha \}.$$

Therefore for any $\theta \in \Theta_0$,

$$ P_\theta (p \leq u) = P_\theta (\inf \{ \alpha : \{X_1, \ldots, X_n\} \in R_\alpha \} \leq u) \leq P_\theta (\{X_1, \ldots, X_n\} \in R_p, p \leq u) \leq P_\theta (\{X_1, \ldots, X_n\} \in R_u) \quad (\text{Since, } p \leq u \implies R_p \subseteq R_u) \leq u.$$

(b) If there exists a $\theta \in \Theta_0$ such that,

$$ P_\theta (\{X_1, \ldots, X_n\} \in R_\alpha) = \alpha,$$

for every $0 \leq \alpha \leq 1$, then plugging this into the last step in the previous set of equations we get,

$$ P_\theta (p \leq u) \leq P_\theta (\{X_1, \ldots, X_n\} \in R_u) = u.$$

Problem 4 (15 points)

The $FWER$ in this case is given by,

$$ FWER = P \left[ \bigcup_{i=1}^{d} \text{ reject } H_{0i} \right] = 1 - P \left[ \bigcap_{i=1}^{d} \text{ accept } H_{0i} \right] = 1 - P (p_1 \geq \alpha) \leq 1 - 1 + \alpha = \alpha.$$

Notice that we never used the independence of the p-values. Hence this procedure controls the FWER at $\alpha$ even if the p-values are dependent. This procedure has lower power than Bonferroni because if $H_{01}$ is true, then with probability less than or equals $\alpha$, it will be able to reject any of the other hypothesis even if they are false. Hence it has very low power to reject the other hypotheses.

Problem 5 (15 points)

(a) The FWER in this case is given by,

$$ FWER = P \left[ \bigcup_{i=1}^{d} \text{ reject } H_{0i} \right] \leq \sum_{i=1}^{d} P (\text{ reject } H_{0i}) \leq \sum_{i=1}^{d} \frac{\alpha}{2} \frac{1 - \left(\frac{1}{2}\right)^d}{\frac{1}{2}} = \alpha \left(1 - \left(\frac{1}{2}\right)^d\right) \leq \alpha.$$

Hence her procedure controls the FWER at level $\alpha$. 

(b) Since FWER $\geq$ FDR, in this case,

$$FDR \leq FWER \leq \alpha.$$ 

Hence, FDR is also controlled at level $\alpha$. 