שאלה 1

$$P(Y=1|X=1)=\Phi(\beta_0+\beta_1 X),\qquad \hat{\beta}=\begin{pmatrix}0.7803\\-0.5253\end{pmatrix}, Cov(\hat{\beta})=\begin{pmatrix}0.0145 & -0.0092\\-0.0092 & 0.0101\end{pmatrix}=V$$

$$\hat{\theta}=x^T\hat{\beta}\sim N(\theta,x^TVx)\rightarrow\begin{cases}\theta_L=\hat{\theta}-z_{1-\frac{\alpha}{2}}\sqrt{x^TVx}\\ \theta_U=\hat{\theta}+z_{1-\frac{\alpha}{2}}\sqrt{x^TVx}\end{cases}\rightarrow P(Y=1|X=1)\in[\Phi(\theta_L),\Phi(\theta_U)]$$

$$x=\begin{pmatrix}1\\1\end{pmatrix}\rightarrow\hat{\theta}=0.7803-0.5253=0.255$$

$$x^TVx=(1,1)\begin{pmatrix}0.0145 & -0.0092\\-0.0092 & 0.0101\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}=(1,1)\begin{pmatrix}0.0053\\0.0009\end{pmatrix}=0.0062\rightarrow\sqrt{x^TVx}=0.0787$$

$$\alpha=0.05\rightarrow z_{1-\frac{\alpha}{2}}=1.96\rightarrow z_{1-\frac{\alpha}{2}}\sqrt{x^TVx}=0.1543$$

$$\begin{cases}\theta_L=0.255-0.1543=0.1007\\ \theta_U=0.255+0.1543=0.4093\end{cases}\rightarrow P(Y=1|X=1)\in[\Phi(0.1007),\Phi(0.4093)]$$

$$\rightarrow P(Y=1|X=1)\in[0.5401,0.6589]$$

בשל מגבלת הערכים בטבלה שמצורפת למבחן, התקבל גם הפתרון $[0.5398,0.6591]$, לפי $[\Phi(0.10),\Phi(0.41)]$.

שאלה 2

נסמן $Q_m\sim Pois\left(e^{\alpha+\theta_m}\right), m=1,\dots,M$ וכן $\sum_m Q_m=N$, $\sum_m e^{\theta_m}=1$, כמו כן נסמן $\beta_m=e^{\alpha+\theta_m}$. הסכום של מ"מ פואסוניים $x_1\sim Pois(\beta_1),\dots,x_M\sim Pois(\beta_M)$ מתפלג פואסון עם פרמטר $\sum_m\beta_m$ ולכן:

$$\sum_m\beta_m=\sum_m e^{\alpha+\theta_m}=\sum_m e^{\alpha}\cdot e^{\theta_m}=e^{\alpha}\sum_m e^{\theta_m}=e^{\alpha}\rightarrow P(N=n)=\frac{(\sum_m\beta_m)^n e^{-\sum_m\beta_m}}{n!}=\frac{(e^{\alpha})^n e^{-e^{\alpha}}}{n!}$$

מכאן כי הפילוג המותנה יהיה:

$$P(Q_1,\dots,Q_M|N=n)=\frac{\prod_{m=1}^M\frac{\left(e^{\alpha+\theta_m}\right)^{Q_m}e^{-e^{\alpha+\theta_m}}}{Q_m!}}{\frac{e^{n\alpha}e^{-e^{\alpha}}}{n!}}=\frac{n!}{\prod_{m=1}^M Q_m!}\cdot\frac{\prod_{m=1}^M\left(e^{\alpha+\theta_m}\right)^{Q_m}e^{-e^{\alpha+\theta_m}}}{(e^{\alpha})^n e^{-e^{\alpha}}}$$

$$=\frac{n!}{\prod_{m=1}^M Q_m!}\cdot\frac{\prod_{m=1}^M(e^{\alpha})^{Q_m}\prod_{m=1}^M\left(e^{\theta_m}\right)^{Q_m}\prod_{m=1}^M e^{-e^{\alpha+\theta_m}}}{(e^{\alpha})^n e^{-e^{\alpha}}}$$

$$=\frac{n!}{\prod_{m=1}^M Q_m!}\cdot\frac{(e^{\alpha})^{\sum_m Q_m}\prod_{m=1}^M\left(e^{\theta_m}\right)^{Q_m}e^{-\sum_m e^{\alpha+\theta_m}}}{(e^{\alpha})^n e^{-e^{\alpha}}}=\frac{n!}{\prod_{m=1}^M Q_m!}\cdot\frac{(e^{\alpha})^n\prod_{m=1}^M\left(e^{\theta_m}\right)^{Q_m}e^{-e^{\alpha}\overbrace{\sum_m e^{\theta_m}}^{=1}}}{(e^{\alpha})^n e^{-e^{\alpha}}}$$

$$=\frac{n!}{\prod_{m=1}^M Q_m!}\cdot\frac{(e^{\alpha})^n\prod_{m=1}^M\left(e^{\theta_m}\right)^{Q_m}e^{-e^{\alpha}}}{(e^{\alpha})^n e^{-e^{\alpha}}}=\frac{n!}{\prod_{m=1}^M Q_m!}\cdot\prod_{m=1}^M\left(e^{\theta_m}\right)^{Q_m}=Mult\left(n;e^{\theta_1},\dots,e^{\theta_M}\right)$$

בלומר הפילוג המותנה הוא מולטינומי.

שאלה 3

א. נשתמש בפרמטריזציה מהצורה $Y = W\eta + \epsilon$:

עבור המודל המלא

$$\begin{pmatrix}92\\80\\65\\75\\86\\91\\56\\69\\56\\34\\57\\54\end{pmatrix} = Y = \begin{pmatrix}1&1&0&1\\1&1&0&1\\1&1&0&-1\\1&1&0&-1\\1&0&1&1\\1&0&1&1\\1&0&1&-1\\1&0&1&-1\\1&-1&-1&1\\1&-1&-1&1\\1&-1&-1&-1\\1&-1&-1&-1\end{pmatrix}\begin{pmatrix}\mu\\\alpha_1\\\alpha_2\\\beta_1\end{pmatrix} + \epsilon, (X^TX) = \begin{pmatrix}12&0&0&0\\0&8&4&0\\0&4&8&0\\0&0&0&12\end{pmatrix} \to (X^TX)^{-1} = \frac{1}{12}\begin{pmatrix}1&0&0&0\\0&2&-1&0\\0&-1&2&0\\0&0&0&1\end{pmatrix}$$

$$X^TY = \begin{pmatrix}815\\111\\101\\63\end{pmatrix} \to \hat{\eta} = (X^TX)^{-1}X^TY = \frac{1}{12}\begin{pmatrix}815\\121\\91\\63\end{pmatrix} = \begin{pmatrix}67.9167\\10.0833\\7.5833\\5.25\end{pmatrix}$$

$$\hat{Y} = \begin{pmatrix}83.25\\83.25\\72.75\\72.75\\80.75\\80.75\\70.25\\70.25\\55.5\\55.5\\45\\45\end{pmatrix} \to e = Y - \hat{Y} = \begin{pmatrix}8.75\\-3.25\\-7.75\\2.25\\5.25\\10.25\\-14.25\\-1.25\\0.5\\-21.5\\12\\9\end{pmatrix} \to SSE = \|e\|^2 = 1177$$

עבור המודל החלקי (פתרון מלא; ביוון שהמודל מאוזן, ניתן היה פשוט להשמיט את האומד של β_1 תוך מתן הסבר מתאים)

$$\begin{pmatrix}92\\80\\65\\75\\86\\91\\56\\69\\56\\34\\57\\54\end{pmatrix} = Y = \begin{pmatrix}1&1&0\\1&1&0\\1&1&0\\1&1&0\\1&0&1\\1&0&1\\1&0&1\\1&0&1\\1&-1&-1\\1&-1&-1\\1&-1&-1\\1&-1&-1\end{pmatrix}\begin{pmatrix}\mu\\\alpha_1\\\alpha_2\end{pmatrix} + \epsilon, \left(X^{0^T}X^0\right) = \begin{pmatrix}12&0&0\\0&8&4\\0&4&8\end{pmatrix} \to \left(X^{0^T}X^0\right)^{-1} = \frac{1}{12}\begin{pmatrix}1&0&0\\0&2&-1\\0&-1&2\end{pmatrix}$$

$$X^{0^T}Y = \begin{pmatrix}815\\111\\101\end{pmatrix} \to \hat{\eta}^0 = \left(X^{0^T}X^0\right)^{-1}X^{0^T}Y = \frac{1}{12}\begin{pmatrix}815\\121\\91\end{pmatrix} = \begin{pmatrix}67.9167\\10.0833\\7.5833\end{pmatrix}$$

$$\hat{Y}^0=\begin{pmatrix}78\\78\\78\\78\\75.5\\75.5\\75.5\\75.5\\50.25\\50.25\\50.25\\50.25\end{pmatrix}\rightarrow e^0=\hat{Y}-\hat{Y}^0=\begin{pmatrix}5.25\\5.25\\-5.25\\-5.25\\5.25\\5.25\\-5.25\\-5.25\\5.25\\5.25\\-5.25\\-5.25\end{pmatrix}\rightarrow SSE^0-SSE=\|e^0\|^2=330.75$$

$$\rightarrow F=\frac{\frac{SSE^0}{1}}{\frac{SSE}{8}}=2.2481\sim\mathcal{F}_{1,8}$$

פתרון נוסף לסעיף זה: כיוון שמדובר במודל מאוזן, ניתן לפתור גם באמצעות הנוסחא

$$F_{AB}=\frac{\frac{n\sum_i\sum_j\left(\bar{y}_{ij.}-\bar{y}_{i..}-\bar{y}_{.j.}+\bar{y}_{...}\right)^2}{(I-1)(J-1)}}{\frac{\sum_i\sum_j\sum_k\left(y_{ijk}-\bar{y}_{ij.}\right)^2}{N-IJ}}$$

ב. תחת H_0 יש 9 דרגות חופש, יש שני קונטרסטים לכן הערך הקריטי יהיה $t_{1-\frac{\omega}{2K},9}=t_{1-\frac{0.05}{4},9}=t_{0.9875,9}$. נמצא ערך זה ע"י אינטרפולציה:

$$t_{0.9875,9}=t_{0.975,9}+\frac{0.9875-0.975}{0.99-0.975}\left(t_{0.99,9}-t_{0.975,9}\right)=2.262+\frac{0.0125}{0.015}(2.821-2.262)=2.262+\frac{5}{6}\cdot0.559=2.7278$$

נביט במטריצת הקונטרסטים:

$$\Psi=\begin{pmatrix}1&0&-1\\1&-1&0\end{pmatrix}\begin{pmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{pmatrix}\rightarrow Var(\hat{\Psi})=\sigma^2\begin{pmatrix}\frac{1^2}{4}+\frac{0}{4}+\frac{(-1)^2}{4}\\\frac{1^2}{4}+\frac{(-1)^2}{4}+\frac{0}{4}\end{pmatrix}=\sigma^2\begin{pmatrix}\frac{1}{2}\\\frac{1}{2}\end{pmatrix}$$

האומד לשונות תחת H_0 הוא $s^2=\frac{1}{12-3}\sum\left(Y-\hat{Y}^0\right)^2=\frac{1}{9}\cdot1507.75=167.5278$ ולכן אומד סטיית התקן הוא $s=12.9433$.

האומדים לערכי הקונטרסטים:

$$\Psi=\begin{pmatrix}1&0&-1\\1&-1&0\end{pmatrix}\begin{pmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{pmatrix}=\begin{pmatrix}2&1\\1&-1\end{pmatrix}\begin{pmatrix}\alpha_1\\\alpha_2\end{pmatrix}\rightarrow\hat{\Psi}=\begin{pmatrix}2\hat{\alpha}_1+\hat{\alpha}_2\\\hat{\alpha}_1-\hat{\alpha}_2\end{pmatrix}=\begin{pmatrix}27.75\\2.5\end{pmatrix}$$

רווחי הסמך בשיטת בונפרוני: $\psi^{(k)}\in\left[\hat{\psi}^{(k)}\pm t_{1-\frac{\omega}{2K},df}s\sqrt{V(c)}\right]$

$$(\alpha_1-\alpha_3)\in[27.75\pm2.7278\cdot12.9433\cdot0.7071]=[2.7845,52.7155]$$

$$(\alpha_1-\alpha_2)\in[2.5\pm2.7278\cdot12.9433\cdot0.7071]=[-22.4655,27.4655]$$

פתרון תוך שימוש בערך t הקרוב ביותר מהטבלה ($t_{0.99,9}=2.821$) גם כן התקבל:

$$(\alpha_1-\alpha_3)\in[27.75\pm2.821\cdot12.9433\cdot0.7071]=53.5685$$

$$(\alpha_1-\alpha_2)\in[2.5\pm2.821\cdot12.9433\cdot0.7071]=[-23.3185,28.3185]$$

שאלה 4

א.

$$\hat{F}_n(t) = \frac{1}{n}\sum_{j=1}^{n} I_{\{t_j \le t\}}$$

ב.

$$\hat{F}_n(t) = \begin{cases} 0 & t < 21 \\ 0.1 & 21 \le t < 34 \\ 0.2 & 34 \le t < 45 \\ 0.3 & 45 \le t < 48 \\ 0.4 & 48 \le t < 49 \\ 0.5 & 49 \le t < 50 \\ 0.6 & 50 \le t < 51 \\ 0.7 & 51 \le t < 55 \\ 0.8 & 55 \le t < 73 \\ 0.9 & 73 \le t < 90 \\ 1 & 90 \le t \end{cases}$$

שאלה 5

אם המודל (AB, AC) מתקיים אז: $P(B=j|A=i)P(C=k|A=i) = P(B=j, C=k|A=i) = \frac{P(A=i,B=j,C=k)}{P(A=i)}$

$$\rightarrow P(A=i, B=j, C=k) = P(A=i)P(B=j|A=i)P(C=k|A=i)$$

נבחר כעת את הנתונים הבאים, עבור $I = J = K = 2$:

$$P(A=1) = \frac{1}{3} \rightarrow P(A=2) = \frac{2}{3}$$

$$P(B=1|A=1) = 0.6 \rightarrow P(B=2|A=1) = 0.4$$

$$P(C=1|A=1) = 0.7 \rightarrow P(C=2|A=1) = 0.3$$

בדי שלא יתקיים המודל (AB, C) נדרוש $P(A=i, B=j, C=k) \neq P(A=i, B=j)P(C=k)$, לבן נשנה את ההסתברויות המותנות עבור $A = 2$:

$$P(B=1|A=2) = 0.8 \rightarrow P(B=2|A=2) = 0.2$$

$$P(C=1|A=2) = 0.6 \rightarrow P(C=2|A=2) = 0.4$$

לפי הנוסחא $P(A=i, B=j, C=k) = P(A=i)P(B=j|A=i)P(C=k|A=i)$ נקבל את הנתונים הבאים:

A	B	C	π
1	1	1	0.14
1	1	2	0.06
1	2	1	0.09333
1	2	2	0.04
2	1	1	0.32
2	1	2	0.21333
2	2	1	0.08
2	2	2	0.05333

נכפיל את ההסתברויות ב-300 כדי לקבל מספרי תצפיות שלמים:

A	B	C	π	f
1	1	1	0.14	42
1	1	2	0.06	18
1	2	1	0.09333	28
1	2	2	0.04	12
2	1	1	0.32	96
2	1	2	0.21333	64
2	2	1	0.08	24
2	2	2	0.05333	16

בעת נותר להוכיח כי המודל (AB, C) אינו מתקיים. נביט בהסתברות למאורע $(A = 1, B = 1)$:

$$f_{111} = 0.14, f_{112} = 0.06 \rightarrow P(A = 1, B = 1) = 0.2$$

$$f_{111} = 0.14, f_{121} = 0.09333, f_{211} = 0.32, f_{221} = 0.08 \rightarrow P(C = 1) = 0.6333$$

$$P(A = 1, B = 1) \cdot P(C = 1) = 0.2 \cdot 0.6333 = 0.12667 \neq 0.14 = P(A = 1, B = 1, C = 1)$$

כלומר $P(A = i, B = j, C = k) \neq P(A = i, B = j) \cdot P(C = k)$ וגמרנו.