Άσκηση 3. Τα παρακάτω δεδομένα δίνουν το πλήθος y_{jk} των φυτών που διατήρησαν μια ιδιότητα όταν n_{jk} φυτά βρέθηκαν κάτω από διαφορετικές συνθήκες.

Ένας ποιοτικός παράγοντας (αγωγή) ήταν η αποθήκευση σε θερμοκρασία 3^0C για 48 ώρες ή η αποθήκευση σε κανονικές συνθήκες. Μία άλλη συμμεταβλητή ήταν η χρησιμοποίηση φυγόκεντρης δύναμης σε τρεις διαφορετικές τιμές 40, 150, 350.

Συνθήκη αποθήκευσης		*Φυγόκεντρος* 40	150	350
Κανονική	y_{1k}	55	52	57
	n_{1k}	102	99	108
Αγωγή	y_{2k}	55	50	50
	n_{2k}	76	81	90

Πάρε ως συμμεταβλητές τις $x_1 = \log 40 = 3.689$, $x_2 = \log 150 = 5.011$ και $x_3 = \log 350 = 5.858$

α) Κάνε γραφική παράσταση των αναλογιών $p_{jk} = y_{jk} / n_{jk}$ ως προς x_k.

β) Εξέτασε τα τρία μοντέλα (εκτίμησε παραμέτρους, αποκλίσεις, εκτιμημένες αναλογίες):

i) logit $\pi_{jk} = \alpha_j + \beta_j x_k$, $j = 1$ κανονικές συνθήκες, $j = 2$ αγωγή.

ii) logit $\pi_{jk} = \alpha_j + \beta\, x_k$,

iii) logit $\pi_{jk} = \alpha + \beta\, x_k$,

γ) Κάνε στο μοντέλο (i) τον έλεγχο $\beta_1 = \beta_2$ σε στάθμη 0.05. Επίσης κάνε τον έλεγχο $\alpha_1 = \alpha_2$ στο μοντέλο (ii) σε στάθμη 0.05.

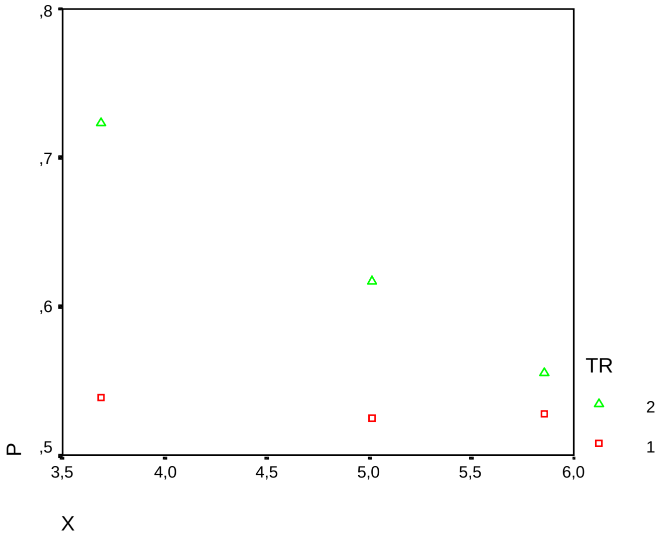
Απαντήσεις Άσκησης 3. Θεωρούμε ότι οι μεταβλητές απόκρισης $Y_{11},Y_{12},Y_{13},Y_{21},Y_{22},Y_{23}$ (πλήθος φυτών που διατήρησαν την ιδιότητα σε κάθε μία από τις 6 περιπτώσεις) προέρχονται από διωνυμική κατανομή με παραμέτρους (m_{jk}, π_{jk}). Τα π_{jk} μάλιστα εξαρτώνται από τις συνθήκες ως εξής (έχουμε τρία μοντέλα):

$$\text{i) } \pi_{jk} = \frac{e^{a_j+\beta_j x_k}}{1+e^{a_j+\beta_j x_k}}, \quad \text{ii) } \pi_{jk} = \frac{e^{a_j+\beta x_k}}{1+e^{a_j+\beta x_k}}, \quad \text{iii) } \pi_{jk} = \frac{e^{a+\beta x_k}}{1+e^{a+\beta x_k}}$$

Αρχικά εισάγονται τα δεδομένα στο SPSS ως εξής ($p = y/n$):

y	*n*	*x*	*tr*	*p*
55	102	3.689	1	0,5392
52	99	5.011	1	0,5253
57	108	5.858	1	0,5278
55	76	3.689	2	0,7237
50	81	5.011	2	0,6173
50	90	5.858	2	0,5556

α) Η γραφική παράσταση των αναλογιών $p_{jk} = y_{jk} / n_{jk}$ ως προς x_k γίνεται χρησιμοποιώντας τη διαδικασία Graphs/Scatter/Simple, Yaxis:p, Xaxis:x, Set Markers by : tr



Παρατηρούμε ότι στην περίπτωση που η αποθήκευση γίνεται σε θερμοκρασία 3^0C (αγωγή, tr = 2) η αύξηση της φυγόκεντρης δύναμης οδηγεί σε μείωση του ποσοστού των φυτών που διατήρησαν την ιδιότητα που εξετάζουμε. Αντίθετα, στην άλλη περίπτωση (αποθήκευση σε κανονικές συνθήκες, tr = 1) παρατηρούμε ότι η αύξηση της φυγόκεντρης δύναμης δεν φαίνεται να οδηγεί σε αύξηση ή σε μείωση του ίδιου ποσοστού. Αυτό αποτελεί μία πρώτη ένδειξη για το ότι δεν πρέπει το μοντέλο logit $\pi_{jk} = \alpha + \beta\, x_k$ να είναι το καλύτερο (διότι στο μοντέλο αυτό δεν περιλαμβάνεται η επίδραση της θερμοκρασίας αποθήκευσης).

β) i) Το πρώτο μοντέλο logit $\pi_{jk} = \alpha_j + \beta_j\, x_k$ γράφεται αναλυτικά για τις 6 παρατηρήσεις ως εξής:

$$
\begin{aligned}
\text{logit } \pi_{11} &= \alpha_1 + \beta_1\, x_1 \\
\text{logit } \pi_{12} &= \alpha_1 + \beta_1\, x_2 \\
\text{logit } \pi_{13} &= \alpha_1 + \beta_1\, x_3 \\
\text{logit } \pi_{21} &= \qquad\qquad \alpha_2 + \beta_2\, x_1 \\
\text{logit } \pi_{22} &= \qquad\qquad \alpha_2 + \beta_2\, x_2 \\
\text{logit } \pi_{23} &= \qquad\qquad \alpha_2 + \beta_2\, x_3
\end{aligned}
$$

Επειδή η διαδικασία του SPSS που θα χρησιμοποιήσουμε προυποθέτει μοντέλο με σταθερά, θα το μετασχηματίσουμε έτσι ώστε να έχει σταθερά. Το παραπάνω μοντέλο γράφεται ισοδύναμα:

$$
\begin{aligned}
\text{logit } \pi_{11} &= \alpha_1 + \beta_1 x_1 \\
\text{logit } \pi_{12} &= \alpha_1 + \beta_1 x_2 \\
\text{logit } \pi_{13} &= \alpha_1 + \beta_1 x_3 \\
\text{logit } \pi_{21} &= \alpha_1 + \beta_1 x_1 + \gamma + \delta\, x_1 \\
\text{logit } \pi_{22} &= \alpha_1 + \beta_1 x_2 + \gamma + \delta\, x_2 \\
\text{logit } \pi_{23} &= \alpha_1 + \beta_1 x_3 + \gamma + \delta\, x_3
\end{aligned}
$$

όπου $\gamma = \alpha_2 - \alpha_1$, $\delta = \beta_2 - \beta_1$ ή με μορφή πινάκων:

$$
\text{logit}\begin{bmatrix} \pi_{11} \\ \pi_{12} \\ \pi_{13} \\ \pi_{21} \\ \pi_{22} \\ \pi_{23} \end{bmatrix} = \begin{bmatrix} 1 & x_1 & 0 & 0 \\ 1 & x_2 & 0 & 0 \\ 1 & x_3 & 0 & 0 \\ 1 & x_1 & 1 & x_1 \\ 1 & x_2 & 1 & x_2 \\ 1 & x_3 & 1 & x_3 \end{bmatrix} \begin{bmatrix} a_1 \\ \beta_1 \\ \gamma \\ \delta \end{bmatrix}
$$

Άρα προσθέτουμε δύο ακόμη στήλες (z_1, z_2) στα δεδομένα:

y	n	x	z_1	z_2	tr	p
55	102	3.689	0	0	1	0,5392
52	99	5.011	0	0	1	0,5253
57	108	5.858	0	0	1	0,5278
55	76	3.689	1	3.689	2	0,7237
50	81	5.011	1	5.011	2	0,6173
50	90	5.858	1	5.858	2	0,5556

και το πρώτο μοντέλο γράφεται: logit $\pi = \alpha_1 + \beta_1 x + \gamma z_1 + \delta z_2$. Χρησιμοποιούμε τη διαδικασία Analyze / regression / probit με Response: y, Total: m, Covariates: x, z1, z2, Model logit από όπου προκύπτει ο εξής πίνακας:

```
* * * * * * * * * * * *  P R O B I T    A N A L Y S I S  * * * * * * * * * * * *
DATA  Information
          6 unweighted cases accepted.
          0 cases rejected because of missing data.
          6 cases are in the control group.
MODEL Information
      ONLY Logistic Model is requested.
* * * * * * * * * * * *  P R O B I T    A N A L Y S I S  * * * * * * * * * * * *
 Parameter estimates converged after 12 iterations.
 Optimal solution found.
 Parameter Estimates (LOGIT model:  (LOG(p/(1-p))) = Intercept + BX):

           Regression Coeff.  Standard Error     Coeff./S.E.
   X               -,02275          ,12685         -,17937
   Z1              1,97708          ,99808         1,98089
   Z2              -,31860          ,19888        -1,60199

                   Intercept  Standard Error  Intercept/S.E.
                      ,23397          ,62842          ,37231

  Pearson  Goodness-of-Fit  Chi Square =       ,028    DF = 2   P =  ,986
 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
 Covariance(below) and Correlation(above) Matrices of Parameter Estimates
                  X          Z1          Z2
 X           ,01609     ,61918    -,63782
 Z1          ,07839     ,99616    -,98442
 Z2         -,01609    -,19540     ,03955
```

Σύμφωνα με τα αποτελέσματα της παραπάνω διαδικασίας, οι εκτιμήσεις των παραμέτρων μετά από 12 βήματα της επαναληπτικής διαδικασίας είναι

$$\hat{a}_1 = 0.23397,\ \hat{\beta}_1 = -0.02275,\ \hat{\gamma} = 1.97708,\ \hat{\delta} = 0.31860,$$

και άρα

$$\hat{a}_2 = \hat{a}_1 + \hat{\gamma} = 0.23397 + 1.97708 = 2.21105,\ \hat{\beta}_2 = \hat{\beta}_1 + \hat{\delta} = -0.02275 - 0.31860 = -0.34135$$

ενώ το χι-τετράγωνο του Pearson είναι ίσο με **0.028** (με $n-p = 6-4 = 2$ β.ε.). Το p-value που αφορά τον έλεγχο της υπόθεσης H_0: $g(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta}$ είναι ίσο με

$$p-value = P(X^2 > 0.028 \mid X^2 \sim \chi_2^2) \approx 0.986.$$

Σύμφωνα λοιπόν με το χι-τετράγωνο του Pearson, δεν έχουμε αρκετά στοιχεία ώστε να απορρίψουμε ότι το μοντέλο είναι σωστό (δηλ. το μοντέλο γίνεται αποδεκτό).

```
* * * * * * * * * * * *  P R O B I T    A N A L Y S I S  * * * * * * * * * * * *
 Observed and Expected Frequencies

              Number of    Observed    Expected
          X    Subjects   Responses   Responses    Residual      Prob
       3,69       102,0        55,0      54,819        ,181    ,53744
       5,01        99,0        52,0      52,465       -,465    ,52995
       5,86       108,0        57,0      56,716        ,284    ,52515
       3,69        76,0        55,0      54,832        ,168    ,72147
       5,01        81,0        50,0      50,429       -,429    ,62258
       5,86        90,0        50,0      49,738        ,262    ,55265
```

Οι τρεις πρώτες στήλες του παραπάνω πίνακα είναι οι x, m, y αντίστοιχα. Η 4η στήλη (expected responses) περιέχει τις εκτιμημένες (fitted) τιμές $\hat{\mu}_{jk} = \hat{y}_{jk} = m_{jk}\hat{\pi}_{jk}$ (πρόκειται για τον εκτιμημένο αναμενόμενο αριθμό φυτών που διατήρησαν την ιδιότητα σε κάθε μία από τις 3 τιμές τις φυγόκεντρης δύναμης x_k και για $j = 1$ (κανονικές συνθήκες), 2 (αγωγή)). Η στήλη με τα κατάλοιπα περιέχει τις διαφορές $y_{jk} - \hat{y}_{jk}$. Τέλος, η στήλη Prob αποτελείται από τα $\hat{\pi}_{jk}$ (εκτιμημένες αναλογίες)

$$\hat{\pi}_{1k} = \frac{e^{\hat{a}_1+\hat{\beta}_1 x_k}}{1+e^{\hat{a}_1+\hat{\beta}_1 x_k}}\ (tr=1),\ \ \hat{\pi}_{2k} = \frac{e^{(\hat{a}_1+\hat{\gamma})+(\hat{\beta}_1+\hat{\delta})x_k}}{1+e^{(\hat{a}_1+\gamma)+(\hat{\beta}_1+\hat{\delta})x_k}}\ (tr=2),\ \ k=1,2,3.$$

Επίσης η απόκλιση (deviance) μπορεί να υπολογιστεί μέσω των καταλοίπων απόκλισης όπως και στις Ασκήσεις 1,2. Υπενθυμίζεται ότι

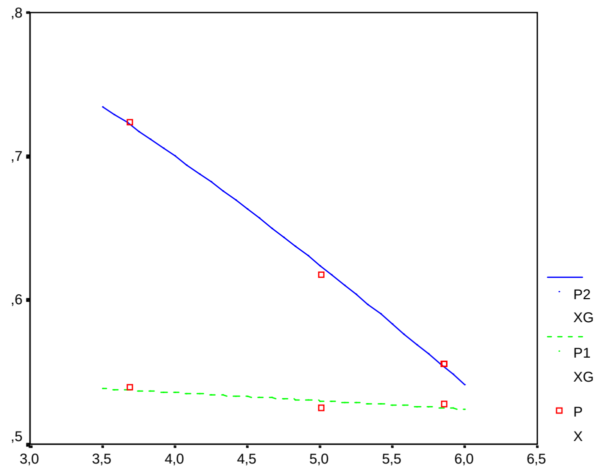
$$r_i^D = \mathrm{sgn}(y_i - m_i\,\hat{\pi}_i)\sqrt{2}\sqrt{y_i \ln\frac{y_i}{m_i\,\hat{\pi}_i} + (m_i - y_i)\ln\frac{m_i - y_i}{m_i - m_i\,\hat{\pi}_i}}$$

και επομένως κατασκευάζουμε (compute) μία νέα στήλη με τα κατάλοιπα απόκρισης r_i^D

```
r_d = (y-n*ep)/ABS(y-n*ep)*(2**0.5)*ABS(y*LN(y/(n*ep))
        +(n-y)*LN((n-y)/(n-n*ep)))**0.5
```

όπου ep είναι οι εκτιμημένες αναλογίες. Επειδή ως γνωστό, $D_{c,f} = \sum_{i=1}^{n}(r_i^D)^2$, κατασκευάζουμε τη στήλη **r_d2 = r_d**2** και τέλος (Analyze/Descriptive Statistics/Descriptives/r_d2, options: sum) προκύπτει ότι η απόκλιση είναι ίση με D = 0.02764.

Όπως και σε παραπάνω ασκήσεις, η γραφική παράσταση $\hat{\pi}(x)$ μαζί με τα $(x_i, y_i/m_i)$ θα είναι (έχουμε δύο καμπύλες, μία για $tr = 1$ (κανονικές συνθήκες) και μία για $tr = 2$ (αγωγή) με διαφορετικό intercept και slope):



δηλαδή,

$$\hat{\pi}_1(x) = \frac{e^{\hat{a}_1+\hat{\beta}_1 x}}{1+e^{\hat{a}_1+\hat{\beta}_1 x}}\ (tr=1),\quad \hat{\pi}_2(x) = \frac{e^{(\hat{a}_1+\hat{\gamma})+(\hat{\beta}_1+\hat{\delta})x}}{1+e^{(\hat{a}_1+\hat{\gamma})+(\hat{\beta}_1+\hat{\delta})x}}\ (tr=2).$$

ii) Το δεύτερο μοντέλο logit $\pi_{jk} = \alpha_j + \beta x_k$ γράφεται αναλυτικά για τις 6 παρατηρήσεις:

$$\begin{aligned}
\text{logit } \pi_{11} &= \alpha_1 + \beta x_1 \\
\text{logit } \pi_{12} &= \alpha_1 + \beta x_2 \\
\text{logit } \pi_{13} &= \alpha_1 + \beta x_3 \\
\text{logit } \pi_{21} &= \qquad \beta x_1 + \alpha_2 \\
\text{logit } \pi_{22} &= \qquad \beta x_2 + \alpha_2 \\
\text{logit } \pi_{23} &= \qquad \beta x_3 + \alpha_2
\end{aligned}$$

ή ισοδύναμα:

$$\begin{aligned}
\text{logit}\ \pi_{11} &= \alpha_1 + \beta x_1 \\
\text{logit}\ \pi_{12} &= \alpha_1 + \beta x_2 \\
\text{logit}\ \pi_{13} &= \alpha_1 + \beta x_3 \\
\text{logit}\ \pi_{21} &= \alpha_1 + \beta x_1 + \gamma \\
\text{logit}\ \pi_{22} &= \alpha_1 + \beta x_2 + \gamma \\
\text{logit}\ \pi_{23} &= \alpha_1 + \beta x_3 + \gamma
\end{aligned}$$

όπου $\gamma = \alpha_2 - \alpha_1$ και άρα το συγκεκριμένο μοντέλο γράφεται: logit $\pi = \alpha_1 + \beta x + \gamma z_1$. Στη συνέχεια χρησιμοποιούμε τη διαδικασία Analyze / regression / probit με Response: y, Total: m, Covariates: x, z1, Model logit από όπου προκύπτει ο εξής πίνακας:

```
* * * * * * * * * * * *  P R O B I T    A N A L Y S I S  * * * * * * * * * * * *
DATA  Information
          6 unweighted cases accepted.
          0 cases rejected because of missing data.
          3 cases are in the control group.
MODEL Information
      ONLY Logistic Model is requested.

* * * * * * * * * * * *  P R O B I T    A N A L Y S I S  * * * * * * * * * * * *

 Parameter estimates converged after 9 iterations.
 Optimal solution found.

 Parameter Estimates (LOGIT model:  (LOG(p/(1-p))) = Intercept + BX):

           Regression Coeff.  Standard Error     Coeff./S.E.

   X                -,15460          ,09703        -1,59335
   Z1                ,40684          ,17462         2,32981

                   Intercept  Standard Error  Intercept/S.E.
                     ,87677          ,48704         1,80022

  Pearson  Goodness-of-Fit  Chi Square =      2,598    DF = 3   P =  ,458

  Since Goodness-of-Fit Chi square is NOT significant, no heterogeneity
  factor is used in the calculation of confidence limits.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
 Covariance(below) and Correlation(above) Matrices of Parameter Estimates

                    X         Z1
 X             ,00941    -,03647
 Z1           -,00062     ,03049
```

Οι εκτιμήσεις των παραμέτρων μετά από 9 βήματα της επαναληπτικής διαδικασίας είναι

$$\hat{a}_1 = 0.87677,\ \hat{\beta} = -0.1546,\ \hat{\gamma} = 0.40684$$

και άρα

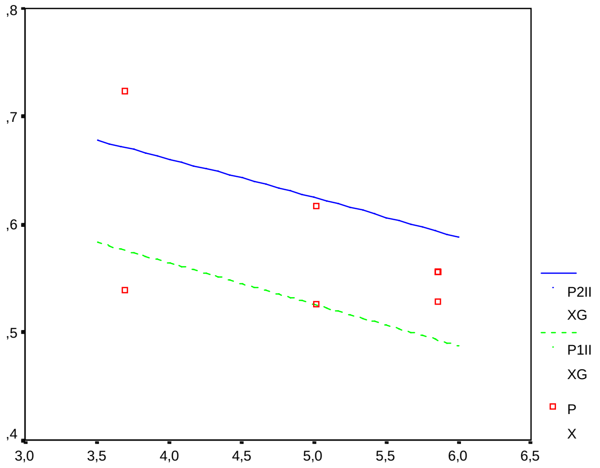
$$\hat{a}_2 = \hat{a}_1 + \hat{\gamma} = 0.87677 + 0.40684 = 1.28361$$

ενώ επίσης το χι-τετράγωνο του Pearson είναι ίσο με **2,598** (με $n-p = 6-3 = 3$ β.ε.). Το p-value που αφορά τον έλεγχο της υπόθεσης H_0: $g(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta}$ είναι ίσο με 0.458 και άρα δεν έχουμε αρκετά στοιχεία ώστε να απορρίψουμε ότι το μοντέλο είναι σωστό (δηλ. το μοντέλο γίνεται αποδεκτό).

```
* * * * * * * * * * * *  P R O B I T    A N A L Y S I S  * * * * * * * * * * * *
 Observed and Expected Frequencies

              Number of    Observed    Expected
         X     Subjects   Responses   Responses     Residual     Prob
      3,69       102,0        55,0      58,754       -3,754   ,57602
      5,01        99,0        52,0      52,025        -,025   ,52550
      5,86       108,0        57,0      53,221        3,779   ,49279
      3,69        76,0        55,0      51,006        3,994   ,67113
      5,01        81,0        50,0      50,589        -,589   ,62456
      5,86        90,0        50,0      53,405       -3,405   ,59339
```

Η στήλη Prob αποτελείται από τις εκτιμημένες αναλογίες $\hat{\pi}_{jk}$. Η απόκλιση (deviance) υπολογίζεται όπως και παραπάνω και βρίσκεται ότι είναι ίση με D = 2.619. Η γραφική παράσταση $\hat{\pi}(x)$ μαζί με τα $(x_i, y_i/m_i)$ είναι (και εδώ έχουμε δύο καμπύλες, μία για $tr = 1$ (κανονικές συνθήκες) και μία για $tr = 2$ (αγωγή) με διαφορετικό intercept και ίδιο slope).



δηλαδή,

$$\hat{\pi}_1(x) = \frac{e^{\hat{a}_1+\hat{\beta}x}}{1+e^{\hat{a}_1+\hat{\beta}x}}\ (tr=1), \quad \hat{\pi}_2(x) = \frac{e^{(\hat{a}_1+\hat{\gamma})+\hat{\beta}x}}{1+e^{(\hat{a}_1+\hat{\gamma})+\hat{\beta}x}}\ (tr=2)\,.$$

β) ii) Το τρίτο μοντέλο logit $\pi_{jk} = \alpha + \beta x_k$ γράφεται αναλυτικά για τις 6 παρατηρήσεις:

$$\text{logit}\ \pi_{11} = \alpha + \beta x_1$$
$$\text{logit}\ \pi_{12} = \alpha + \beta x_2$$
$$\text{logit}\ \pi_{13} = \alpha + \beta x_3$$
$$\text{logit}\ \pi_{21} = \alpha + \beta x_1$$
$$\text{logit}\ \pi_{22} = \alpha + \beta x_2$$
$$\text{logit}\ \pi_{23} = \alpha + \beta x_3$$

και άρα το συγκεκριμένο μοντέλο γράφεται απλά: logit $\pi = \alpha + \beta x$. Χρησιμοποιούμε τη διαδικασία Analyze / regression / probit με Response: y, Total: m, Covariates: x, Model logit από όπου προκύπτει ο πίνακας:

```
* * * * * * * * * * * *  P R O B I T   A N A L Y S I S  * * * * * * * * * * * *
DATA  Information
          6 unweighted cases accepted.
          0 cases rejected because of missing data.
          0 cases are in the control group.
MODEL Information
      ONLY Logistic Model is requested.

* * * * * * * * * * * *  P R O B I T   A N A L Y S I S  * * * * * * * * * * * *

 Parameter estimates converged after 7 iterations.
 Optimal solution found.

 Parameter Estimates (LOGIT model:  (LOG(p/(1-p))) = Intercept + BX):

          Regression Coeff.  Standard Error     Coeff./S.E.
   X               -,14785          ,09650        -1,53213

                  Intercept  Standard Error  Intercept/S.E.
                    1,02136          ,48134         2,12189

  Pearson  Goodness-of-Fit  Chi Square =      7,971    DF = 4    P =  ,093
```

Οι εκτιμήσεις των παραμέτρων μετά από 7 βήματα της επαναληπτικής διαδικασίας είναι

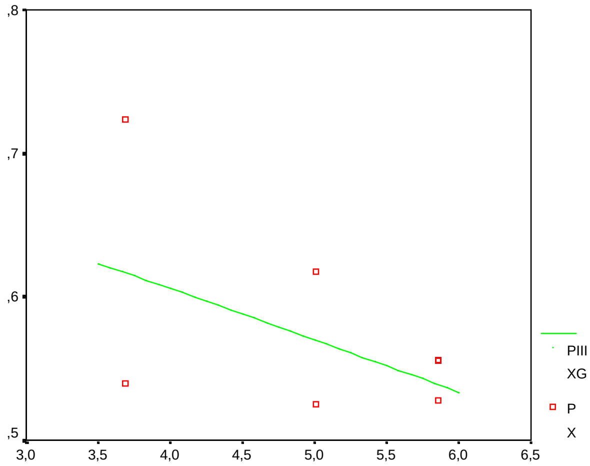
$$\hat{a} = 1.0213,\ \hat{\beta} = -0.14785$$

ενώ επίσης το χι-τετράγωνο του Pearson είναι ίσο με **7,971** (με $n-p = 6-2 = 4$ β.ε.). Το p-value που αφορά τον έλεγχο της υπόθεσης H_0: $g(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta}$ είναι ίσο με 0.093 και άρα δεν έχουμε αρκετά στοιχεία ώστε να απορρίψουμε ότι το μοντέλο είναι σωστό (δηλ. το μοντέλο γίνεται αποδεκτό).

```
* * * * * * * * * * * *  P R O B I T    A N A L Y S I S  * * * * * * * * * * * *
 Observed and Expected Frequencies

              Number of   Observed    Expected
         X    Subjects   Responses   Responses    Residual     Prob
      3,69       102,0        55,0      62,913      -7,913   ,61679
      5,01        99,0        52,0      56,397      -4,397   ,56967
      5,86       108,0        57,0      58,184      -1,184   ,53874
      3,69        76,0        55,0      46,876       8,124   ,61679
      5,01        81,0        50,0      46,143       3,857   ,56967
      5,86        90,0        50,0      48,487       1,513   ,53874
```

Η στήλη Prob αποτελείται από τις εκτιμημένες αναλογίες $\hat{\pi}_{jk}$. Η απόκλιση (deviance) υπολογίζεται όπως και παραπάνω και βρίσκεται ότι είναι ίση με D = 8.091. Η γραφική παράσταση $\hat{\pi}(x)$ μαζί με τα $(x_i, y_i/m_i)$ είναι (εδώ έχουμε μία καμπύλη και για $tr = 1$ (κανονικές συνθήκες) και για $tr = 2$ (αγωγή)).



δηλαδή,

$$\hat{\pi}_j(x) = \frac{e^{\hat{a}+\hat{\beta}x}}{1+e^{\hat{a}+\hat{\beta}x}} \quad j = 1,2\,.$$

γ) i) Ο έλεγχος της υπόθεσης H_0: $\beta_1 = \beta_2$ στο μοντέλο logit $\pi_{jk} = \alpha_j + \beta_j x_k$ είναι ισοδύναμος:

A) με τον έλεγχο H_0: $\delta = 0$. Εδώ απορρίπτουμε την H_0 αν

$$|\hat{\delta}/s(\hat{\delta})| = 1{,}60199 > Z_{a/2} = 1.96$$

άρα *δεχόμαστε* ότι $\delta = 0 \Leftrightarrow \beta_1 = \beta_2$ (ισοδύναμα μπορούμε να ελέγξουμε αν το 0 ανήκει στο δ.ε. 95% για το δ: (–0.708, 0.071)).

B) με τον έλεγχο H_0: logit $\pi_{jk} = \alpha_j + \beta x_k$ έναντι του μεγαλύτερου μοντέλου H_1: logit $\pi_{jk} = \alpha_j + \beta_j x_k$. Εδώ απορρίπτουμε την H_0 (ότι δηλαδή το μικρότερο μοντέλο είναι σωστό) όταν

$$X^2_{Pearson} - X'^2_{Pearson} = 2.598 - 0.028 = 2.57 > \chi^2_{p'-p;a} = \chi^2_{3-2;a} = \chi^2_{1;0.05} = 3.84$$

άρα *δεχόμαστε* ότι το μικρότερο μοντέλο είναι σωστό, δηλαδή ότι $\delta = 0 \Leftrightarrow \beta_1 = \beta_2$ (ο έλεγχος αυτός μπορεί να γίνει και μέσω των διαφορών των αποκλισεων)

ii) Ο έλεγχος της υπόθεσης H_0: $\alpha_1 = \alpha_2$ στο μοντέλο logit $\pi_{jk} = \alpha_j + \beta x_k$ είναι ισοδύναμος:

A) με τον έλεγχο H_0: $\gamma = 0$. Εδώ απορρίπτουμε την H_0 αν

$$|\hat{\gamma} / s(\hat{\gamma})| = 2{,}32981 > Z_{a/2} = 1.96$$

άρα *απορρίπτουμε* ότι γ = 0 ⇔ $\alpha_1 = \alpha_2$. (ισοδύναμα μπορούμε να ελέγξουμε αν το 0 ανήκει στο δ.ε. 95% για το γ: (0.064, 0.750)).

B) με τον έλεγχο H_0: logit $\pi_{jk} = \alpha + \beta x_k$ έναντι του μεγαλύτερου μοντέλου H_1: logit $\pi_{jk} = \alpha_j + \beta x_k$. Εδώ απορρίπτουμε την H_0 (ότι δηλαδή το μικρότερο μοντέλο είναι σωστό) όταν

$$X^2_{Pearson} - X'^2_{Pearson} = 7.971 - 2.598 = 5.373 > \chi^2_{p'-p;a} = \chi^2_{3-2;a} = \chi^2_{1;0.05} = 3.84$$

άρα *απορρίπτουμε* ότι το μικρότερο μοντέλο είναι σωστό, δηλαδή ότι γ = 0 ⇔ $\alpha_1 = \alpha_2$. (ο έλεγχος αυτός μπορεί να γίνει και μέσω των διαφορών των αποκλισεων)