Project3

一、資料

| **Obs** | **rank** | **high** | **low** | **total** | **p** |
| --- | --- | --- | --- | --- | --- |
| **1** | 1 | 273 | 147 | 420 | 0.65 |
| **2** | 2 | 206 | 458 | 664 | 0.310240964 |
| **3** | 3 | 36 | 540 | 576 | 0.0625 |

rank=1:坪數<5

rank=2:坪數5~10坪

rank=3:坪數>10坪

high: number of每坪月租金>=1500

low: number of每坪月租金<1500

二、model of binomial

(1) x:坪數(rank) x=1,2,3





x=1 =0.6244

x=2 =0.3413

x=3 =0.0581

The chance of high rent decreases as rank level increases.

se()=0.012

| **模型資訊** | |
| --- | --- |
| **資料集** | WORK.DATA |
| **分布** | Binomial |
| **連結函數** | Identity |
| **回應變數 (事件)** | high |
| **回應變數 (試驗)** | total |

|  |  |
| --- | --- |
| **讀取的觀測值數目** | 3 |
| **使用的觀測值數目** | 3 |
| **事件數目** | 515 |
| **試驗數目** | 1660 |

| **回應概況** | | |
| --- | --- | --- |
| **已排序的 值** | **二元結果** | **總次數** |
| **1** | 事件 | 515 |
| **2** | 非事件 | 1145 |

| **參數資訊** | |
| --- | --- |
| **參數** | **效果** |
| **Prm1** | Intercept |
| **Prm2** | rank |

| **最大概度參數估計值的分析** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **參數** | **自由度** | **估計值** | **標準 誤差** | **Wald 95% 信賴界限** | | **Wald 卡方** | **Pr > ChiSq** |
| **Intercept** | 1 | 0.9075 | 0.0315 | 0.8458 | 0.9693 | 829.47 | <.0001 |
| **rank** | 1 | -0.2831 | 0.0117 | -0.3061 | -0.2602 | 583.44 | <.0001 |
| **縮放** | 0 | 1.0000 | 0.0000 | 1.0000 | 1.0000 |  |  |

| **觀測值統計值** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **觀測值** | **high** | **total** | **預測值** | **線性預測量** | **線性預測量的標準誤差** | **HessWgt** |
| **1** | **273** | **420** | 0.6244079 | 0.6244079 | 0.0203843 | 1742.2461 |
| **2** | **206** | **664** | 0.3412752 | 0.3412752 | 0.0106243 | 2824.2125 |
| **3** | **36** | **576** | 0.0581426 | 0.0581426 | 0.0092213 | 11257.861 |

(2) x:坪數(rank) x=1,2,3





x=1 =0.6676

x=2 =0.288

x=3 =0.0753

The chance of high rent decreases as rank level increases.

| **模型資訊** | |
| --- | --- |
| **資料集** | WORK.DATA |
| **分布** | Binomial |
| **連結函數** | Logit |
| **回應變數 (事件)** | high |
| **回應變數 (試驗)** | total |

|  |  |
| --- | --- |
| **讀取的觀測值數目** | 3 |
| **使用的觀測值數目** | 3 |
| **事件數目** | 515 |
| **試驗數目** | 1660 |

| **回應概況** | | |
| --- | --- | --- |
| **已排序的 值** | **二元結果** | **總次數** |
| **1** | 事件 | 515 |
| **2** | 非事件 | 1145 |

| **參數資訊** | |
| --- | --- |
| **參數** | **效果** |
| **Prm1** | Intercept |
| **Prm2** | rank |

| **最大概度參數估計值的分析** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **參數** | **自由度** | **估計值** | **標準 誤差** | **Wald 95% 信賴界限** | | **Wald 卡方** | **Pr > ChiSq** |
| **Intercept** | 1 | 2.2998 | 0.1737 | 1.9593 | 2.6402 | 175.31 | <.0001 |
| **rank** | 1 | -1.6025 | 0.0902 | -1.7793 | -1.4256 | 315.51 | <.0001 |
| **縮放** | 0 | 1.0000 | 0.0000 | 1.0000 | 1.0000 |  |  |

| **觀測值統計值** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **觀測值** | **high** | **total** | **預測值** | **線性預測量** | **線性預測量的標準誤差** | **HessWgt** |
| **1** | **273** | **420** | 0.6675887 | 0.6972991 | 0.0946524 | 93.203896 |
| **2** | **206** | **664** | 0.2879902 | -0.905165 | 0.0634565 | 136.15443 |
| **3** | **36** | **576** | 0.0753251 | -2.507629 | 0.1239851 | 40.119095 |

(3) x:坪數(rank) x=1,2,3





x=1 =0.6610

x=2 =0.2965

x=3 =0.0689

The chance of high rent decreases as rank level increases.

| **模型資訊** | |
| --- | --- |
| **資料集** | WORK.DATA |
| **分布** | Binomial |
| **連結函數** | Probit |
| **回應變數 (事件)** | high |
| **回應變數 (試驗)** | total |

|  |  |
| --- | --- |
| **讀取的觀測值數目** | 3 |
| **使用的觀測值數目** | 3 |
| **事件數目** | 515 |
| **試驗數目** | 1660 |

| **回應概況** | | |
| --- | --- | --- |
| **已排序的 值** | **二元結果** | **總次數** |
| **1** | 事件 | 515 |
| **2** | 非事件 | 1145 |

| **參數資訊** | |
| --- | --- |
| **參數** | **效果** |
| **Prm1** | Intercept |
| **Prm2** | rank |

| **最大概度參數估計值的分析** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **參數** | **自由度** | **估計值** | **標準 誤差** | **Wald 95% 信賴界限** | | **Wald 卡方** | **Pr > ChiSq** |
| **Intercept** | 1 | 1.3651 | 0.1009 | 1.1674 | 1.5628 | 183.11 | <.0001 |
| **rank** | 1 | -0.9498 | 0.0503 | -1.0483 | -0.8513 | 357.10 | <.0001 |
| **縮放** | 0 | 1.0000 | 0.0000 | 1.0000 | 1.0000 |  |  |

| **觀測值統計值** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **觀測值** | **high** | **total** | **預測值** | **線性預測量** | **線性預測量的標準誤差** | **HessWgt** |
| **1** | **273** | **420** | 0.6610366 | 0.4152939 | 0.0566891 | 251.90514 |
| **2** | **206** | **664** | 0.2964866 | -0.534532 | 0.0360922 | 382.85915 |
| **3** | **36** | **576** | 0.0688569 | -1.484359 | 0.0666664 | 155.64125 |

三、 poisson loglinear model

 x:坪數(rank) x=1,2,3



 



:每增加一個level of 坪數, decreasing the estimated mean number of high rent

| **模型資訊** | |
| --- | --- |
| **資料集** | WORK.DATA |
| **分布** | Poisson |
| **連結函數** | Log |
| **應變數** | high |

|  |  |
| --- | --- |
| **讀取的觀測值數目** | 3 |
| **使用的觀測值數目** | 3 |

| **最大概度參數估計值的分析** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **參數** | **自由度** | **估計值** | **標準 誤差** | **Wald 95% 信賴界限** | | **Wald 卡方** | **Pr > ChiSq** |
| **Intercept** | 1 | 6.4722 | 0.1047 | 6.2670 | 6.6775 | 3819.66 | <.0001 |
| **rank** | 1 | -0.7539 | 0.0617 | -0.8748 | -0.6330 | 149.32 | <.0001 |
| **縮放** | 0 | 1.0000 | 0.0000 | 1.0000 | 1.0000 |  |  |

四、 negative binomial regression

 x:坪數(rank) x=1,2,3



 



:每增加一個level of 坪數, decreasing the estimated mean number of high rent

| **模型資訊** | |
| --- | --- |
| **資料集** | WORK.DATA |
| **分布** | Negative Binomial |
| **連結函數** | Log |
| **應變數** | high |

|  |  |
| --- | --- |
| **讀取的觀測值數目** | 3 |
| **使用的觀測值數目** | 3 |

| **最大概度參數估計值的分析** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **參數** | **自由度** | **估計值** | **標準 誤差** | **Wald 95% 信賴界限** | | **Wald 卡方** | **Pr > ChiSq** |
| **Intercept** | 1 | 6.8860 | 0.6077 | 5.6950 | 8.0771 | 128.41 | <.0001 |
| **rank** | 1 | -0.9885 | 0.2911 | -1.5589 | -0.4180 | 11.53 | 0.0007 |
| **散佈** | 1 | 0.1165 | 0.1008 | 0.0214 | 0.6351 |  |  |

五、 model checking

1. 檢驗模型

(1) Linear probability model



(2) Logistic regression model



2.檢驗參數

前提假設: 為真

 

Test 

(1)Wald test



Reject 

坪數影響每坪月租金為顯著

(2)the likelihood-ratio test

>3.84

Reject 

坪數影響每坪月租金為顯著

(3) 信賴區間





| **模型資訊** | |
| --- | --- |
| **資料集** | WORK.DATA |
| **分布** | Binomial |
| **連結函數** | Identity |
| **回應變數 (事件)** | high |
| **回應變數 (試驗)** | total |

|  |  |
| --- | --- |
| **讀取的觀測值數目** | 3 |
| **使用的觀測值數目** | 3 |
| **事件數目** | 515 |
| **試驗數目** | 1660 |

| **回應概況** | | |
| --- | --- | --- |
| **已排序的 值** | **二元結果** | **總次數** |
| **1** | 事件 | 515 |
| **2** | 非事件 | 1145 |

| **參數資訊** | |
| --- | --- |
| **參數** | **效果** |
| **Prm1** | Intercept |
| **Prm2** | rank |

| **用於評估配適度的準則** | | | |
| --- | --- | --- | --- |
| **準則** | **自由度** | **值** | **值/自由度** |
| **偏差** | 1 | 4.2688 | 4.2688 |
| **縮放偏差** | 1 | 4.2688 | 4.2688 |
| **Pearson 卡方** | 1 | 4.2174 | 4.2174 |
| **縮放 Pearson X2** | 1 | 4.2174 | 4.2174 |
| **對數概度** |  | -819.9368 |  |
| **完整對數概度** |  | -11.4122 |  |
| **AIC (越小越好)** |  | 26.8245 |  |
| **AICC (越小越好)** |  | . |  |
| **BIC (越小越好)** |  | 25.0217 |  |

| **最大概度參數估計值的分析** | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **參數** | **自由度** | **估計值** | **標準 誤差** | **Wald 95% 信賴界限** | | **概度比 95% 信賴界限** | | **Wald 卡方** | **Pr > ChiSq** |
| **Intercept** | 1 | 0.9075 | 0.0315 | 0.8458 | 0.9693 | 0.8453 | 0.9687 | 829.47 | <.0001 |
| **rank** | 1 | -0.2831 | 0.0117 | -0.3061 | -0.2602 | -0.3057 | -0.2597 | 583.44 | <.0001 |
| **縮放** | 0 | 1.0000 | 0.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |  |  |

| **型 3 分析的 LR 統計值** | | | |
| --- | --- | --- | --- |
| **來源** | **自由度** | **卡方** | **Pr > ChiSq** |
| **rank** | 1 | 416.18 | <.0001 |

| **模型資訊** | |
| --- | --- |
| **資料集** | WORK.DATA |
| **分布** | Binomial |
| **連結函數** | Logit |
| **回應變數 (事件)** | high |
| **回應變數 (試驗)** | total |

|  |  |
| --- | --- |
| **讀取的觀測值數目** | 3 |
| **使用的觀測值數目** | 3 |
| **事件數目** | 515 |
| **試驗數目** | 1660 |

| **回應概況** | | |
| --- | --- | --- |
| **已排序的 值** | **二元結果** | **總次數** |
| **1** | 事件 | 515 |
| **2** | 非事件 | 1145 |

| **參數資訊** | |
| --- | --- |
| **參數** | **效果** |
| **Prm1** | Intercept |
| **Prm2** | rank |

| **用於評估配適度的準則** | | | |
| --- | --- | --- | --- |
| **準則** | **自由度** | **值** | **值/自由度** |
| **偏差** | 1 | 3.5982 | 3.5982 |
| **縮放偏差** | 1 | 3.5982 | 3.5982 |
| **Pearson 卡方** | 1 | 3.5490 | 3.5490 |
| **縮放 Pearson X2** | 1 | 3.5490 | 3.5490 |
| **對數概度** |  | -819.6015 |  |
| **完整對數概度** |  | -11.0769 |  |
| **AIC (越小越好)** |  | 26.1539 |  |
| **AICC (越小越好)** |  | . |  |
| **BIC (越小越好)** |  | 24.3511 |  |

| **最大概度參數估計值的分析** | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **參數** | **自由度** | **估計值** | **標準 誤差** | **Wald 95% 信賴界限** | | **概度比 95% 信賴界限** | | **Wald 卡方** | **Pr > ChiSq** |
| **Intercept** | 1 | 2.2998 | 0.1737 | 1.9593 | 2.6402 | 1.9635 | 2.6448 | 175.31 | <.0001 |
| **rank** | 1 | -1.6025 | 0.0902 | -1.7793 | -1.4256 | -1.7826 | -1.4287 | 315.51 | <.0001 |
| **縮放** | 0 | 1.0000 | 0.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |  |  |

| **型 3 分析的 LR 統計值** | | | |
| --- | --- | --- | --- |
| **來源** | **自由度** | **卡方** | **Pr > ChiSq** |
| **rank** | 1 | 416.85 | <.0001 |
| **模型資訊** | |
| **資料集** | WORK.DATA |
| **分布** | Binomial |
| **連結函數** | Logit |
| **回應變數 (事件)** | high |
| **回應變數 (試驗)** | total |

|  |  |
| --- | --- |
| **讀取的觀測值數目** | 3 |
| **使用的觀測值數目** | 3 |
| **事件數目** | 515 |
| **試驗數目** | 1660 |

| **回應概況** | | |
| --- | --- | --- |
| **已排序的 值** | **二元結果** | **總次數** |
| **1** | 事件 | 515 |
| **2** | 非事件 | 1145 |

| **用於評估配適度的準則** | | | |
| --- | --- | --- | --- |
| **準則** | **自由度** | **值** | **值/自由度** |
| **偏差** | 2 | 420.4491 | 210.2246 |
| **縮放偏差** | 2 | 420.4491 | 210.2246 |
| **Pearson 卡方** | 2 | 391.7704 | 195.8852 |
| **縮放 Pearson X2** | 2 | 391.7704 | 195.8852 |
| **對數概度** |  | -1028.0269 |  |
| **完整對數概度** |  | -219.5024 |  |
| **AIC (越小越好)** |  | 441.0048 |  |
| **AICC (越小越好)** |  | 445.0048 |  |
| **BIC (越小越好)** |  | 440.1034 |  |

| **最大概度參數估計值的分析** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **參數** | **自由度** | **估計值** | **標準 誤差** | **Wald 95% 信賴界限** | | **Wald 卡方** | **Pr > ChiSq** |
| **Intercept** | 1 | -0.7990 | 0.0531 | -0.9030 | -0.6950 | 226.77 | <.0001 |
| **縮放** | 0 | 1.0000 | 0.0000 | 1.0000 | 1.0000 |  |  |

六、程式碼

**data** data;

input rank high low total;

cards;

1 273 147 420

2 206 458 664

3 36 540 576

;

**proc** **print** data=data;

**run**;

**PROC** **GENMOD**;

MODEL high/total=rank / DIST=BIN LINK=IDENTITY P;

TITLE 'BINOMAL';

**PROC** **GENMOD**;

MODEL high/total=rank / DIST=BIN LINK=LOGIT P;

TITLE 'LOGIT';

**PROC** **GENMOD**;

MODEL high/total=rank / DIST=BIN LINK=PROBIT P;

TITLE 'PROBIT';

**PROC** **GENMOD**;

MODEL high=rank/DIST=POI LINK=LOG;

title 'poisson log';

**PROC** **GENMOD**;

MODEL high=rank/DIST=NEGBIN LINK=LOG;

title 'negative log';

**PROC** **GENMOD**; MODEL high/total=rank / DIST=BIN LINK=IDENTITY WALDCI LRCI TYPE3 OBSTATS;

**PROC** **GENMOD**; MODEL high/total=rank / DIST=BIN LINK=LOGIT WALDCI LRCI TYPE3 OBSTATS;

**PROC** **GENMOD**; MODEL high/total=/ DIST=BIN LINK=LOGIT;