**Questions**

1. See the problem statement and equation on page 1 and define 𝜋 for this specific problem. Your answer should start with: *“π in this problem will be the probability of an applicant getting hired by this firm (based on an applicant’s education, experience, and gender). [y=1]*

*P (An applicant Getting hired by education, experience, gender) = 𝜋*

**2.** Define “odds” for this specific problem. You can either write the equation for the odds ratio or answer like: “odds in this problem are the probability of (or that) *an applicant getting hired P(y=1)* **divided** by the probability of (or that) *an applicant not getting hired P(y=0)*.”

**3.** What does 𝜋∗ in the equation on page 1 represent?

𝜋∗= ln ()= 𝛽0 + 𝛽1x1+ 𝛽2x2+ 𝛽3x3

𝜋∗= ln ()= 𝛽0 + 𝛽Education+ 𝛽Experience+ 𝛽Gender

*𝜋∗ (natural logs of odds) represents the probability of an applicant being hired given this model based on education, experience, and gender by an applicant not getting hired by a firm based on education, experience, and gender.*

**4.** What is the Global Null Hypothesis for this specific problem (express that with **both** Betas and Chi-square)? In your SAS printout, use the Likelihood Ratio row in the Global Null Hypothesis section and reason whether this is a statistically useful model or not (use 𝛼=0.05 to reason).

H0: 𝛽1= 𝛽2= 𝛽3= 0

H0: 𝛽Education= 𝛽Experience= 𝛽Gender = 0

**OR**

H0: = 0

H0: = 0

| **Testing Global Null Hypothesis: BETA=0** | | | |
| --- | --- | --- | --- |
| **Test** | **Chi-Square** | **DF** | **Pr > ChiSq** |
| **Likelihood Ratio** | 20.4299 | 3 | 0.0001 |
| **Score** | 15.0320 | 3 | 0.0018 |
| **Wald** | 5.2883 | 3 | 0.1519 |

= 20.4299 and the p-value is 0.0001.

***Where***  *and*  = 14.735

= 20.4299

*At any level of 𝛼, there is enough evidence to indicate that the proposed model (with 3 independent variables) is statistically significant in decreasing the lack of fit.*

| **Model Fit Statistics** | | |
| --- | --- | --- |
| **Criterion** | **Intercept Only** | **Intercept and Covariates** |
| **AIC** | 37.165 | 22.735 |
| **SC** | 38.497 | 28.064 |
| **-2 Log L** | 35.165 | 14.735 |

| **Analysis of Maximum Likelihood Estimates** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Parameter** | **DF** | **Estimate** | **Standard Error** | **Wald Chi-Square** | **Pr > ChiSq** |
| **Intercept** | 1 | -14.2482 | 6.0805 | 5.4909 | 0.0191 |
| **Education** | 1 | 1.1549 | 0.6023 | 3.6767 | 0.0552 |
| **Experience** | 1 | 0.9098 | 0.4293 | 4.4919 | 0.0341 |
| **Gender** | 1 | 5.6037 | 2.6028 | 4.6352 | 0.0313 |

**5.** Is *Years of Higher Education* a statistically useful factor in the above model? Why? Use 𝛼=0.05 to reason.

H0: 𝛽Education = 0

*H0:  = 0*

*– = 0*

*= 3.6767*

*p-value = 0.0552> 0.05*

*The p-value should be smaller than 0.05(level of 𝛼). We assumed that βeducation=0 (the null hypothesis) SAS calculated a p-value of 0.0552, which corroborates that the p-value is greater than 0.05. Thus, we cannot reject the null hypothesis and conclude that there is enough evidence to indicate that the Years of Higher education are not statistically useful in our proposed model of getting hired by the firm*

*Given the 𝛼 = 0.05, there is enough evidence to state that Years of Higher education is not a statistically significant variable in contributing to the probability of getting hired by the firm in our proposed model.*

**6.** Is *Years of Higher Experience* a statistically useful factor in the above model? Why? Use 𝛼=0.05 to reason.

*H0: 𝛽Experience = 0*

*H0:  = 0*

*– = 0*

*= 4.4919*

*P-value = 0.0341< 0.05*

*The p-value should be smaller than 0.05(level of 𝛼). We assumed that βexperience=0 (the null hypothesis) SAS calculated a p-value of 0.0341, which corroborates that the p-value is less than 0.05. Thus, we reject the null hypothesis and conclude that there is enough evidence to indicate that the Years of Higher experience are statistically useful in our proposed model in terms of getting hired by the firm.*

*Given the 𝛼 = 0.05, there is enough evidence to state that Years of Higher Experience is a statistically significant variable in contributing to the probability of getting hired by the firm in our proposed model.*

**7.** Interpret 𝛽2 (the impact of Experience on the odds of Getting Hired) in layman’s terms.

**eexperience = e**0.9098**= 2.484**

**(eexperience – 1) \*100= (2.484 - 1) \*100 = 148%**

| **Odds Ratio Estimates** | | | |
| --- | --- | --- | --- |
| **Effect** | **Point Estimate** | **95% Wald Confidence Limits** | |
| **Education** | 3.174 | 0.975 | 10.333 |
| **Experience** | 2.484 | 1.071 | 5.762 |
| **Gender** | 271.422 | 1.653 | >999.999 |

*Holding other variables fixed, the odds of getting hired increased by 148% for every additional year of experience.*

**8.** What is the Null Hypothesis corresponding to *Gender* in this problem (express that with **both** Betas and Chi-square)?

H0: 𝛽Gender = 0

H0:  = 0

– = 0

**9.** Using the SAS findings, can you reject the Null Hypothesis in Question #8? In other words, does *Gender* statistically matters in getting hired by this particular firm? Why? Use 𝛼=0.05 to reason.

= 4.6352

P-value = 0.0313 < 0.05

*The p-value should be smaller than 0.05(level of 𝛼). We assumed that βgender=0 (the null hypothesis) SAS calculated a p-value of 0.0313, which corroborates that the p-value is less than 0.05. Thus, we reject the null hypothesis and conclude that there is enough evidence to indicate that gender is a statistically useful variable in our proposed model in terms of getting hired by the firm.*

*Given the 𝛼 = 0.05, there is enough evidence to state that Gender is a statistically significant variable in contributing to the probability of getting hired by the firm in our proposed model.*

**10.** *Gender* is a categorical (binary) variable. Suppose that Gender is equal to zero (female applicant). In that case, when Gender is increased by one, it becomes equal to one (male applicant). In other words, gender is changing from Female to Male. Try to use this logic to interpret 𝛽3 (the impact of Gender on getting hired) in layman’s terms. What does this interpretation tell you about discrimination against female applicants in the hiring practices of this particular firm?

**EGender = e**5.6037\*1 =271.422

**(EGender – 1) \* 100 = 27,042\*%**

*Holding other variables fixed, the odds of an applicant being hired increased by 27042% when the applicant is a male in our proposed model.*

**EGender = e**5.6037\*0 =1

**(EGender – 1) \* 100 = 0%**

*Holding other variables fixed, the odds of an applicant being hired do not increase/decrease when the applicant is a female in our proposed model. Hence, we conclude that there may be sex discrimination in the hiring process.*