**Universidade de São Paulo**

**Faculdade de Filosofia, Letras e Ciências Humanas**

**Departamento de Ciência Política**

1º semestre / 2022

Prova Final

**Instruções**

Esta prova representa a avaliação final desta disciplina. Ela deve ser feita individualmente; porém, é esperado que vocês conversem e discutam a respeito destes exercícios. Não será tolerada, sob hipótese alguma, qualquer forma de plágio.

A entrega será feita via *moodle* até as 23:59h do dia 08/01/2023 através do link adequado criado para tal. Não serão recebidas as respostas por nenhum outro meio sob nenhuma justificativa.

Cada discente deverá enviar um arquivo com o próprio nome e número USP tanto no título do arquivo, como na primeira página do arquivo com as respostas. O modelo a ser seguido para o nome do arquivo é *Glauco Peres\_NUSP 1104679*. Aqueles que por ventura não tenham o número USP, utilizem apenas o próprio nome. Serão aceitos arquivos em formato .doc, .docx ou .pdf.

Sejam detalhistas, claros e específicos em suas respostas, mostrando as etapas dos raciocínios. Não apresentem apenas resultados, mas utilizem comentários e discussões envolvidas. Não apresentem saídas diretas dos softwares, se for o caso, mas utilizem tabelas ou gráficos que ajudem na compreensão dos resultados quando for conveniente. Não há uma única maneira de responder cada exercício e, assim, espera-se que vocês desenvolvam as respostas de maneira ampla.

Bom trabalho!

**Enunciados**

**Questão 1**

Considere a seguinte tabela com os resultados de três modelos distintos de regressão. A variável dependente é o total de votos recebido por um candidato em uma eleição para um assento no legislativo estadual no país.

Tabela 1 – Resultados da Regressão – Variável Dependente: Total de votos

|  |  |  |  |
| --- | --- | --- | --- |
|  | (1) | (2) | (3) |
| X1 | 13,65 | 2,80 | 3,20 |
|  | (0,02)\*\*\* | (0,01)\*\*\* | (0,02)\*\*\* |
|  |  |  |  |
| X2 |  | -8.500 | -4.000 |
|  |  | (17,90)\* | (20,81)\*\* |
|  |  |  |  |
| X1 \* X2 |  |  | -1,40 |
|  |  |  | (0,20)\*\*\* |
|  |  |  |  |
| Constante | 700 | 450,0 | 400 |
|  | (2,42)\*\*\* | (0,56)\*\*\* | (2,81)\*\*\* |
| N | 850 | 850 | 850 |
| R2 | 31% | 35% | 87% |
| Erros padrão em parêntesis | | |  |
| \*\*\* p < 0,1%; \*\* p < 1%; \* p < 5% | | |  |

Considere que a primeira variável explicativa (X1) é o valor gasto com a campanha (em milhares de R$) e a segunda variável explicativa (X2) é uma *dummy* que indica se o candidato se declara como uma mulher (assumindo valor igual a 1) ou não.

A partir do que é apresentado nesta tabela, responda às questões a seguir:

1. Escreva a equação que representa o único modelo bivariado estimado. Interprete-a;
2. Qual a diferença de especificação entre os 2º e o 3º modelos? Explique;
3. Interprete o efeito distinto do financiamento de campanha para uma candidata que se declare uma mulher em relação aos demais candidatos;
4. Desenhe em um gráfico as curvas com o efeito do financiamento de campanha sobre os votos para candidatos que se declaram mulheres e para os demais candidatos. Coteje a análise gráfica com sua resposta ao item anterior.

**Questão 2**

Uma analista está avaliando um modelo Poisson para modelar os gols em uma partida de futebol. Para isso, usa uma base de dados com todas as partidas válidas pelo campeonato brasileiro entre 2010 e 2020. Em 4.180 partidas consideradas, a frequência de gols marcados pelos mandantes e pelos visitantes está indicada na tabela abaixo:

Tabela 1 – Tabela descritiva de frequência de gols por situação da equipe

|  |  |  |  |
| --- | --- | --- | --- |
| Gols | Visitante | Mandante | Total |
| 0 | 1.598 | 928 | 2.526 |
| 1 | 1.519 | 1.470 | 2.989 |
| 2 | 760 | 1.047 | 1.807 |
| 3 | 232 | 503 | 735 |
| 4 | 60 | 168 | 228 |
| 5 | 10 | 50 | 60 |
| 6 | 1 | 14 | 15 |
| Total | 4.180 | 4.180 | 8.360 |

Ela constrói esse histograma com os dois resultados, com as respectivas curvas de densidade estimadas:



Em seu modelo inicial, a analista considera como variável dependente o número de gols marcado por um time e como única variável independente uma *dummy* (chamada *home*) que indica se o time é mandante (assumindo valor igual a 1 neste caso) ou não.

Estimando os parâmetros através de um método de Máximo Verossimilhança com uma função logarítmica aplicada à distribuição de Poisson, a analista encontra estes resultados em suas estimativas:

|  |  |  |
| --- | --- | --- |
|  | (1) | (2) |
| VARIÁVEIS | Gols | Gols |
|  |  |  |
| Home |  | 0,4108\*\*\* |
|  |  | (0,02) |
| Constante | 0,190\*\*\* | -0,0363\*\* |
|  | (0,01) | (0,02) |
|  |  |  |
| N | 8.360 | 8.360 |
| Erros padrão em parênteses | |  |
| \*\*\* p<0,01, \*\* p<0,05, \* p<0,1 | |  |

Considerando estes resultados, responda:

1. Mostre que os três valores estimados para os dois modelos acima estão de acordo com a tabela descritiva de frequência de gols por situação da equipe;
2. Encontre a probabilidade de um time mandante fazer 1 gol ou menos em uma partida (*Dica*: você precisará utilizar a função massa de probabilidade de uma distribuição Poisson para isso);
3. Repetindo o método aplicado na questão anterior, encontre a probabilidade de uma partida terminar empatada sem gols.

Na mesma base de dados, a analista encontra que a vitória dos mandantes ocorre em 49,11% dos jogos registrados (2.053 partidas). Pretende assim montar um modelo que capture as chances de um mandante vencer um jogo. Monta duas especificações: em uma, utiliza um modelo não linear (logit) e na segunda, aplica um modelo linear de probabilidade. Introduz quatro variáveis explicativas: o saldo de gols dos times mandante e do time visitante e a posição na tabela dos times mandante e visitante todas mensuradas até o início da partida. A tabela abaixo apresenta os resultados dos modelos:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 1ª especificação | | | 2ª especificação | | |
|  | (1) | (2) | (3) | (4) | (5) | (6) |
| Variáveis | vit | vit | vit | vit | vit | vit |
|  |  |  |  |  |  |  |
| Saldo Gols\_Mand |  | 0.0288\*\*\* | 0.0205\*\*\* |  | 0.00678\*\*\* | 0.00470\*\*\* |
|  |  | (0.00332) | (0.00519) |  | (0.000767) | (0.00122) |
| Saldo Gols\_Vis |  | -0.0256\*\*\* | -0.0190\*\*\* |  | -0.00605\*\*\* | -0.00438\*\*\* |
|  |  | (0.00328) | (0.00514) |  | (0.000767) | (0.00121) |
| Posição\_Mand |  |  | -0.0172\*\* |  |  | -0.00448\*\* |
|  |  |  | (0.00848) |  |  | (0.00204) |
| Posição\_Vis |  |  | 0.0140\* |  |  | 0.00363\* |
|  |  |  | (0.00844) |  |  | (0.00204) |
| Constante | -0.0354 | -0.0247 | 0.00456 | 0.491\*\*\* | 0.494\*\*\* | 0.502\*\*\* |
|  | (0.0309) | (0.0315) | (0.116) | (0.00773) | (0.00761) | (0.0282) |
|  |  |  |  |  |  |  |
| Observações | 4,180 | 4,180 | 4,180 | 4,180 | 4,180 | 4,180 |
| R-quadrado |  |  |  | 0.000 | 0.033 | 0.035 |
| Erros-padrão em parênteses | |  |  |  |  |  |
| \*\*\* p<0.01, \*\* p<0.05, \* p<0.1 | |  |  |  |  |  |

Considerando os resultados apresentados, responda:

1. As especificações conduzem aos mesmos resultados? Discuta;
2. Os sinais dos parâmetros associados às variáveis “Posição\_Mand” e “Posição\_Vis” estão corretos? Interprete-os;
3. Qual a probabilidade estimada para o modelo completo em cada uma das especificações de que o time mandante vença se ele ocupa a vice-liderança na tabela e tenha 5 gols de saldo, ao enfrentar um time que ocupa a 14ª posição na tabela e que tenha 0 gols de saldo?