

SOC3070 Análisis de Datos Categóricos

Trabajo 2 (parte 1)

Información

Ponderación: 20% de la nota final del curso. Entrega: Desde el momento de entrega, los estudiantes tienen 2 semanas exactas de plazo para completar esta tarea. Responder las preguntas *bonus* NO es un requisito necesario para obtener puntaje completo. Responder incorrectamente las preguntas *bonus* no afectará negativamente la nota obtenida, pero responderlas correctamente mejorará la nota obtenida en un máximo de 0.7 puntos (o en la cantidad necesaria para obtener nota máxima si la nota original fuera superior a 6.3)

Introducción

En esta tarea usarán regresión logística multinomial (MNL) y regresión logística ordenada (OLR) para re-analizar un subconjunto de los datos utilizados en el artículo *“It’s not just how the game is played, it’s whether you win or lose”* (2019). Este estudio utiliza un experimento online para identificar el efecto causal de las desigualdades de oportunidades y de resultados sobre creencias acerca de las causas de la desigualdad y percepciones de justicia.

En particular, deberán estimar los efectos de las dos manipulaciones experimentales del estudio sobre las probabilidades de atribuir los resultados de la competencia al talento (“talent”), la suerte (“luck”) o circunstancias externas al individuo (“rules”). Estas respuestas se encuentran en la variable LTC. Una de las manipulaciones experimentales es el status de ganador o perdedor de los participantes (variable W). La otra es el nivel de desigualdad de oportunidades en la competencia, el cual es determinado por el tratamiento al que cada individuo fue asignado (variable T). Para esta tarea usarán el subconjunto de datos correspondientes a los tratamientos RA (“random exchange”, es decir, “level of redistribution” = 0), RE1 (“regressive exchange” con “level of redistribution” = 1) y RE2 (“regressive exchange” con “level of redistribution” = 2).

Para mayor contexto pueden revisar el artículo en el link indicado en el repositorio. Igualmente, los datos están disponibles en el repositorio del curso para ser descargados.

```
data_paper <- read_csv("data_t_2.csv") %>% mutate(W=if_else(W=="TRUE",1,0))
data_paper %>% glimpse()
```

```
## Rows: 660
## Columns: 8
## $ groupId <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2...
## $ playerId <dbl> 2, 1, 1, 2, 1, 2, 2, 1, 1, 2, 2, 1, 2, 1, 1, 2, 2, 1, 2, 1...
## $ fair <chr> "Fair", "Unfair", "Unfair", "Fair", "Fair", "Fair", "Unfair...
## $ T <chr> "RA", "RA", "RA", "RA", "RA", "RA", "RA", "RA", "RA", "RA", "RA"...
## $ W <dbl> 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 1, 1, 0...
## $ LTC <chr> "luck", "luck", "luck", "luck", "luck", "luck", "luck", "luck", "l...
## $ age <dbl> 27, 53, 33, 33, 33, 28, 25, 46, 30, 23, 30, 44, 60, 65, 33...
## $ gender <chr> "Female", "Female", "Male", "Female", "Female", "Male", "M...
```

Ejercicios

1. Usa regresión logística multinomial para estimar el modelo implícito en el panel derecho de la figura 2 (interacción entre variables T y W). Presenta un `summary()` de los resultados. Usa `LTC = "rules"` como categoría de referencia en la variable dependiente.

*Para fijar categoría de referencia puedes usar el siguiente código

```
data_paper$LTC <- factor(data_paper$LTC, levels = c("rules","talent","luck"))
```

- 1.1 Interpreta los coeficientes asociados a TRE1 y TRE2:W.
- 1.2 Transforma e interpreta los coeficientes correspondiente RE1-talent y RE1-luck en términos de odds-ratios (o probabilidades relativas). Presenta el desarrollo formal de una de estas odds-ratios.
- 1.3. Calcula la odds ratio de atribuir los resultados al talento en vez de a la suerte entre los perdedores del tratamiento RA vs los perdedores de RE2.
- 1.4 Manipula los resultados del modelo para obtener las probabilidades esperadas de que los participantes del tratamiento RA – tanto ganadores como perdedores – sostengan que los resultados de la competencia se deben principalmente al “talento”. Expresa formalmente las ecuaciones correspondiente a estas predicciones.
- 1.5 Agrega un control por `age` al modelo estimado en 1. Presenta un `summary()` de los resultados.
- 1.6. (Bonus 1) De acuerdo al modelo estimado en 1.5., ¿cuál es *efecto* (marginal) de edad sobre la probabilidad de que un perdedor de 20 años en el tratamiento RA atribuya los resultados al talento?

2. Usa regresión logística ordinal para estimar el modelo implícito en el panel derecho de la figura 2 (interacción entre variables T y W). Presenta un `summary()` de los resultados. Usa el siguiente orden para la variable LTC: “rules”, “luck”, “talent” (de menos a más “meritocrático”).

*Para el orden de los valores puedes usar el siguiente código

```
data_paper$LTC <- ordered(data_paper$LTC, levels = c("rules","luck","talent"))
```

- 2.1 Interpreta los coeficientes asociados a TRE1 y TRE2:W.
- 2.2 Transforma e interpreta el coeficientes correspondiente a RE1 en términos de odds-ratios acumulativas . Presenta el desarrollo formal de dicha odds-ratio.
- 2.3 Manipula los resultados del modelo para obtener las probabilidades esperadas de que los participantes del tratamiento RA – tanto ganadores como perdedores – sostengan que los resultados de la competencia se deben principalmente al “talento”. Expresa formalmente las ecuaciones correspondiente a estas predicciones.
- 2.4 Agrega un control por `age` al modelo estimado en 2. Presenta un `summary()` de los resultados.
- 2.5. (Bonus 2) De acuerdo al modelo estimado en 2.5., ¿cuál es *efecto* (marginal) de edad sobre la probabilidad de que un perdedor de 20 años en el tratamiento RA atribuya los resultados al talento? Interpreta este resultado.