Ejercicio 1

HENC

2023-10-03

# Instrucciones

1. El ejercicio tiene 2 secciones. Cada una vale 10 puntos:

* Preguntas conceptuales
* Ejercicio práctico

1. Si gusta utilice este archivo de word y conteste en el mismo documento (Pege su código, salidas y gráficas en las secciones relevantes)

# Parte 1: Conceptos

Seleccione o indique la o las opciones correctas.

## Teoría y conceptos

1. **Usted tiene a su cargo evaluar el impacto de un programa piloto para pequeñas y medianas empresas dedicadas a la producción de vasos de vidrio. La intervención consiste en capacitar a las empresas en distintos frentes: lenguaje inglés, procesos modernos de producción, gestión financiera, publicidad y manejo de insumos. El piloto se consideraría exitoso si hay un aumento del 50% de las exportaciones de las empresas al final de la intervención. Como grupo de control se tiene otro grupo de empresas, las cuales reciben el mismo financiamiento, pero no tienen la capacitación. Los grupos se generan aleatoriamente, pero la muestra se considera “pequeña”.**

**Al final de la intervención usted encuentra que el incremento promedio es del 192% con un p-value=.04. Sin embargo, un colega le filtra información de que la capacitación fue heterogénea y que probablemente la variable binaria que utiliza en el modelo de regresión está contaminada por alta variabilidad indeseable. ¿Qué concluiría al respecto?**

1. Dado que en econometría el error de medición reduce el tamaño del efecto, un p-value de .04 me indica que, a pesar del error de medición, existe 96% de probabilidad de que la intervención haya sido exitosa.
2. En econometría el error de medición reduce el tamaño del efecto y un p-value de .04 es indicativo de que la intervención tuvo efectos significativos y muy altos.
3. El poder del piloto está comprometido. En presencia de error de medición el p-value de .04 me dice que probablemente los resultados no son significativos y que el efecto es mucho menor.
4. El efecto de 192% puede estar seriamente sub o sobre estimado debido al alto error de medición. No existe evidencia para inferir que la intervención fue exitosa.
5. **El INEGI, le pide a dos equipos de investigación que generen índice de competitividad económica municipal. Para ello les da una base de datos con 100 variables. El primer equipo utiliza varios métodos de aprendizaje de máquina para tratar de elucidar la relación entre las variables y concluye que hay 10 variables que tienen varianza común. Concluye que el índice debe tener esas variables. El segundo equipo, solamente hace un análisis factorial exploratorio y concluye que hay 11 variables que tienen varianza común. Al contrastar los resultados, el INEGI selecciona los resultados del primer equipo porque provienen de un método más robusto de reducción de variables. El INEGI le pide su opinión:**
6. El INEGI no debería de tomar ninguna de las dos soluciones. Ninguna está tratando de medir la competitividad municipal. El debate se concentra en la superioridad de métodos pero no gira en torno a la medición del fenómeno en cuestión.
7. El INEGI hace bien en elegir la solución que proviene del método más robusto puesto que es la más confiable. Seguro tiene menos error.
8. El análisis factorial exploratorio es adecuado para este tipo de problemas y sus resultados son los más confiables.
9. **Se publican los resultados de una nueva medición de pobreza infantil. Estos resultados fueron validados con correlaciones parciales entre los distintos indicadores -todos correlacionan entre sí-, se agregan con pesos diferenciales para producir puntajes para cada niño en la muestra, y se hace remuestreo -bootstrapping- para calcular los intervalos de confianza de la proporción de niños que viven en pobreza. Otro equipo de investigación calcula Omega, encuentra que el valor total es de .65 y hace un llamado a no utilizar los resultados debido a la baja confiabilidad. ¿Qué diría usted al respecto?**
10. La baja confiabilidad implica baja replicabilidad. Sin embargo, si la técnica de remuestreo muestra que sistemáticamente -con distintas muestras- los resultados se parecen mucho, entonces la medición se replica.
11. El hecho de tomar muestras distintas para calcular los intervalos de confianza de una prevalencia no tiene nada que ver con replicabilidad de los scores de cada persona. Uno puede replicar el valor puntual y tener resultados aleatorios en los scores a nivel individual. La replicabilidad es baja.
12. Se trata de cosas distintas. Una es replicabilidad del estimador puntual y el otro es replicabilidad de los scores individuales. Sin embargo, debido a que lo que se busca es replicar resultados agregados, yo confiaría más en el remuestreo.
13. **Una de las hipótesis más importantes en las nuevas teorías del desarrollo es que los países no se desarrollan por la baja inversión en Innovación. Dos macroeconomistas se proponen a poner a prueba la hipótesis. El primero usa la medición tradicional de cuentas nacionales, que básicamente consiste en tomar la formación de capital en el sector etiquetado como Innovación -Piense que el capital físico y el software no son lo mismo-. El segundo argumenta que la inversión en Innovación es un concepto latente y que requiere definición y medición específica. Así que produce una medición ad hoc. Los dos estiman modelos de series de tiempo con los mismos datos, donde apenas tienen 21 observaciones. Al comparar la de uno y otro trabajo, se encuentra que los efectos son pequeños, casi de la misma magnitud y significativos con p<.05. ¿Qué diría usted al respecto?**
14. Las betas podrán ser iguales pero su significado interpretativo es distinto. En el primer caso, probablemente existe un alto error de medición y, por tanto, la podría bien ser positiva o negativa.
15. Los resultados son los mismos porque la inversión en Innovación se mide casi igual en ambos estudios.
16. Los resultados son los mismos y ambas hipótesis son válidas.
17. **Le piden que evalúe si un test de 50 preguntas puede ser útil para que los médicos diagnostiquen la fragilidad durante la vejez. Una condición que le ponen es que los puntajes se agreguen de manera simple para que los médicos no batallen con la interpretación de los scores. ¿Qué propondría al respecto?**
18. Quedarme con las preguntas que respeten equivalencia Tau. Por lo tanto, elegiría con los ítems que tienen cargas factoriales estadísticamente iguales. Eso garantizaría alta confiabilidad y que las sumas simples lleven a buenos scores.
19. Tirar los ítems que violen equivalencia Tau y que lleven a una alta confiabilidad. Esto es, los ítems cuyos errores sean distintos.
20. Quedarme con los ítems con altas cargas factoriales y que lleven a una alta confiabilidad.
21. Ninguna de las anteriores.
22. **Una escala de 6 indicadores tiene un (). Sin embargo, los valores de las cargas factoriales y los umbrales son exactamente iguales. La escala se quiere utilizar para diagnosticar baja, moderada y alta demencia. Qué concluiría de la escala?**
23. La escala es únicamente útil para el nivel latente equivalente al valor de los umbrales de cada indicador. Por lo tanto, sirve para hacer buenos diagnósticos de niveles medios de demencia.
24. La escala no es útil para hacer distinciones entre grupos. Mismos umbrales significan mismos niveles de severidad.
25. La alta confiabilidad garantiza bajos falsos positivos y negativos, por lo tanto se puede hacer para distinguir entre diferentes grupos.

# Parte 2.

El INEGI levantó una encuesta con un n=10,000 para medir bienestar subjetivo (“Practica\_1”). El INEGI importó una escala de estados unidos que consiste de 15 reactivos que se agregan de manera simple para producir un score que va de 0 a 15. Sin embargo, un grupo de investigadores alega que la escala no es adecuada para México y que no debe implementarse tal y cómo se hizo en EUA. Estos investigadores sugieren, a partir de un modelo teórico aplicado en Latinoamérica, que un modelo de tres dimensiones es más adecuado. Bajo este modelo los primeros cinco indicadores se agregarían en la primera dimensión, los siguientes cinco en la segunda y los últimos cinco en la tercera.

El INEGI te contrata para resolver esta disputa y conteste las siguientes preguntas:

1. ¿Cuál de las dos estructuras propuestas es mejor?
2. ¿Cuál es la confiabilidad de los puntajes de cada propuesta?
3. ¿Respetan, estadísticamente hablando, los indicadores equivalencia Tau?
4. Discuta qué ventajas y desventajas tiene cada propuesta

D<-fread("Practica\_1\_1.dat")