Prueba Parcial – SOLUCIONES
Econometría I: Magíster en Economía
Universidad de Santiago de Chile
Semestre 1, 2017

Nombre de alumna/alumno y firma:	
Pregunta 1 (Máximo 30 puntos)	
Pregunta 2 (Máximo 30 puntos)	

Instrucciones Generales

1. Tiene 75 minutos para responder a esta prueba, dividida en 60 puntos.
2. Póngale nombre a todas las hojas.
3. Sólo el profesor puede responder dudas de enunciado, y sólo en voz alta desde el puesto.
4. Se permite el uso de calculadoras, siempre y cuando no cuenten con dispositivos de comunicación.
5. En caso de copia, se sancionará de acuerdo a lo estipulado en el programa del curso.
6. Las tablas de probabilidad necesarias están incluidas en la prueba. Si no sale la cantidad exacta de grados de libertad necesaria para un cálculo determinado, utilice la cantidad de grados de libertad más cercana.

Buena suerte!!

1. **Estimación (Teoría)** [30 puntos] Se cuenta con una muestra de tamaño N extraída de una distribución uniforme $[0, \theta]$. Las observaciones se denotan como $y_1, y_2, \ldots, y_N$, o en formato vectorial simplemente como $\boldsymbol{y}$. Cada observación es independiente e idénticamente distribuida. La función de densidad de probabilidad del uniforme para cada y_i se escribe como $f(y_i|\theta) = \frac{1}{\theta}$ para $0 \leq y \leq \theta$ (y 0 en cualquier otro punto).

   a) (11 puntos) Le interesa encontrar el estimador de máxima verosimilitud $\widehat{\theta}_{MV}$. Escriba la función de verosimilitud para este estimador. [*Pista:* Se sugiere partir con la forma general de una función de máxima verosimilitud, y por último sustituir para el caso específico.] Explique brevemente en palabras la función de máxima verosimilitud.

   Solución: Partimos escribiendo la fdp conjunta:

$$\begin{aligned} f(y_1, ..., y_N|\theta) &= f(y_1|\theta) \times \cdots \times f(y_N|\theta) \\ &= \prod_{i=1}^{N} f(y_i|\theta) \qquad (1) \end{aligned}$$

   y después lo presentamos como la función de verosimilitud:

$$\mathcal{L}(\theta|y_1, \ldots, y_n) = \prod_{i=1}^{N} f(y_i|\theta). \qquad (2)$$

   Ahora, sabemos que $f(y_i|\theta) = \frac{1}{\theta}$, y que cada observación es independiente, y por lo tanto, podemos reescribir (2) como:

$$\begin{aligned} \mathcal{L}(\theta|y_1, \ldots, y_n) &= \prod_{i=1}^{N} f(y_i|\theta) \\ &= \prod_{i=1}^{N} \frac{1}{\theta} \\ &= \theta^{-N}. \qquad (3) \end{aligned}$$

   Nuestra función de verosimilitud está dada por (3). La función de verosimilitud nos dice cuán probable es que habríamos visto los datos dado el parametro θ.

b) (2 puntos) ¿Por qué podríamos preferir la función de log verosimilitud en vez de la función de verosimilitud para estimar utilizando máxima verosimilitud?

Solución: La funcion de log verosimilitud es, a menudo, más conveniente para maximizar computacionalmente. En vez de trabajar con una serie de probabilidades multiplicadas que resulta en un valor muy pequeño, trabajamos con la sumatoria de una serie de log(probabilidades), que resulta en un valor absoluto más grande. Pero de todas formas, ambas funciones alcanzan su máximo en el mismo punto, y por ende la solución es idéntica.

c) (6 puntos) Ahora, en vez de querer estimar θ utilizando máxima verosimilitud, se quiere esitmar $\widehat{\theta}_{MM}$: un estimador de método de momentos. Explique *en palabras* el procedimiento a seguir, y cómo se podría estimar θ utilizando esta metodologia. Note que no es necesario escribir los momentos a estimar. Asegure de especificar cuál(es) momento(s) central(es) se debe utilizar para estimar.

Solución: Primero necesitamos plantear *momentos poblacionales* que caracterizan nuestro proceso de generador de datos (en base a nuestros supuestos del modelo de probabilidad). Estos momentos son momentos centrals, pero dado que no observamos los momentos poblacionales, segundo tenemos que formar los momentos análogos de la muestra. Una vez que hemos formado los momentos de la muestra, se estima el vector de parametros de método de momentos resolviendo el sistema de ecuaciones. En este caso lo ideal para estimar sería utilizar momentos centrales bajos, como el primer o segundo momento central, ya que es posible que momentos centrales más altos podrían no exisitir.

d) (5 puntos) ¿Se puede estimar el parametro θ utilizando *más* de un momento? Si su respuesta es afirmativa, ¿cómo se puede resolver los momentos?

Solución: Sí, se puede. Siempre cuando se cuenta con tantos (o más) momentos que parametros, se puede estimar mediante método de momentos. El caso en que existen más momentos que parametros se conoce como "sobre-identificación". En un sistema sobre identificado, no se puede hacer que todos los momentos se cumplen con exactitud (a menos que hayan momentos redundantes), entonces lo que se hace es minimizar la siguiente función:

$$Q_N(\boldsymbol{\theta}) = \left[\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{h}(\boldsymbol{w}_i,\boldsymbol{\theta})\right]'\left[\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{h}(\boldsymbol{w}_i,\boldsymbol{\theta})\right].$$

La solución a esta función da el/los parametro/s estimado/s $\widehat{\boldsymbol{\theta}}_{MM}$.

e) (6 puntos) Pensando en los dos estimadores de arriba en forma general, ¿existen casos en que uno sería consistente y otro no? ¿Cuál estimador es más robusto a casos en que los supuestos no se cumplen exactamente?
– Esta pregunta fue difícil. La corrección a esta parte da puntos para respuestas parciales que logran explicar la diferencia de flexiblidad de los dos estimadores en base de sus supuestos. – Solución: Los supuestos para el estimador de máxima verosimilitud son más exigentes de los para el estimador de MM. Es posible que con un supuesto acerca de la distribución de un modelo, que aún cuando el supuesto distribucional no es exactamente correcto, los momentos siguen siendo válidos. Dado que el estimador de MM sólo se supone que se sabe algo acerca de ciertas partes de la distribución, es más robusto a casos cuando los supuestos distribucionales no son exactamente correctos. En este sentido, si se toma un supuesto distribucional que no es correcto, el estimador de MV no será consistente. Pero si los momentos de la distribución aún son como supuesto, el estimador de MM sigues estando consistente.

2. **Estimación (Aplicada)** [30 puntos] Existe una hipótesis que víncula la exposición a níveles elevados de plomo a alzas en tasas de crimen. Suponga que el nivel promedio de crimen en un país es 40 por 100,000 habitantes, y se supone que la distribución de tasas de crimen por ciudad dentro del país sigue una distribución normal. Se quiere estimar la tasa de crimen (μ) *solo* en ciudades que tienen altos niveles de plomo. Imagine que se cuenta con una muestra representativa de tasas de crimen un 64 ciudades (independientes) con altos niveles de plomo. El promedio de la tasa de crimen en la muestra es $\widehat{\mu} = 41.67$ por 100.000 habitantes, y la desviación estándar en la muestra es $\widehat{\sigma} = 8$ por 100.000.

   a) (4 puntos) Calcule un estimador del intervalo de confianza de 95 % para el parametro poblacional μ.
   Solución: Sabemos que podemos formar un estimador de intervalo como:

   $$\hat{\mu} \pm t_{(N-1,\alpha/2)}(\hat{\sigma}/\sqrt{N})$$

   . Aquí lo importante es poder calcular el $t_{(N-1,\alpha/2)}$, que de las tablas de probabilidad (aproximada con $N-1=60$ da un valor de 2.00. Entonces, utilizando los valores conocidos:

   $$\begin{aligned} 41{,}67 \pm 2 \times \frac{8}{\sqrt{64}} \\ 41{,}67 \pm 2, \end{aligned} \tag{4}$$

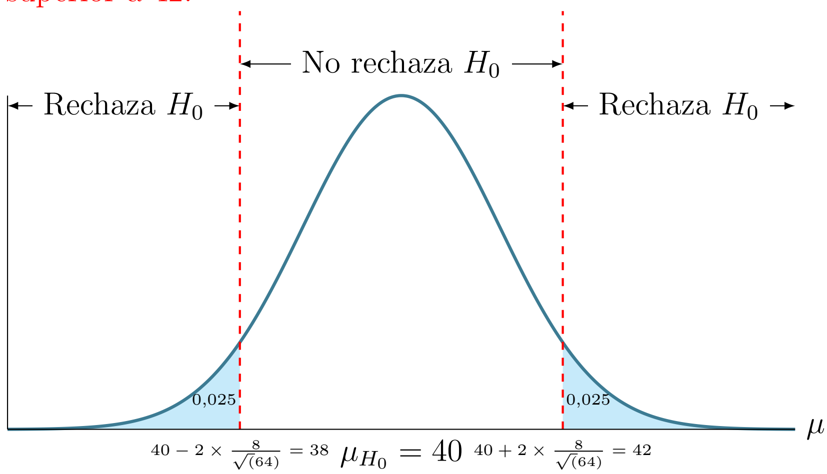
   y el estimador de intervalo es [39,67;43,67].

   b) (5 puntos) Explique en palabras qué nos dice este intervalo de confianza. Asegure de explicar explícitamente qué se puede decir acerca del parametro poblacional con este estimador en mano.
   Solución: Estamos diciendo que si se replicaba el experimento infinitas veces, en 95 % de los casos, nuestro parametro pobacional verdadero va a estar contenido en el estimador del intervalo de confianza. No podemos decir nada acerca del parametro poblacional con un estimador puntual en particular. Dado que el estimador puntual es un valor escalar (desconocido) en un caso particular o está contenido o no está contenido en el estimador

de los intervalos de confianza.

c) (6 puntos) Ahora, imagine que se quiere comprobar la hipótesis nula que $\mu = 40$. Describe la distribución de la tasa de crimen esperado baja la nula que $\mu = 40$. Se sugiere dibujar la distribución para μ, indicando su promedio, y los valores críticos de rechazo a un nivel de significancia de $\alpha = 0{,}05$.

Solución: A continuación se esboza la distribución bajo la nula. Es una distribución t con 63 grados de libertad, pero los valores críticos vienen de la tabla t, entonces $gl = 60 \approx 63$. La región de rechazo es inferior a 38, o superior a 42.



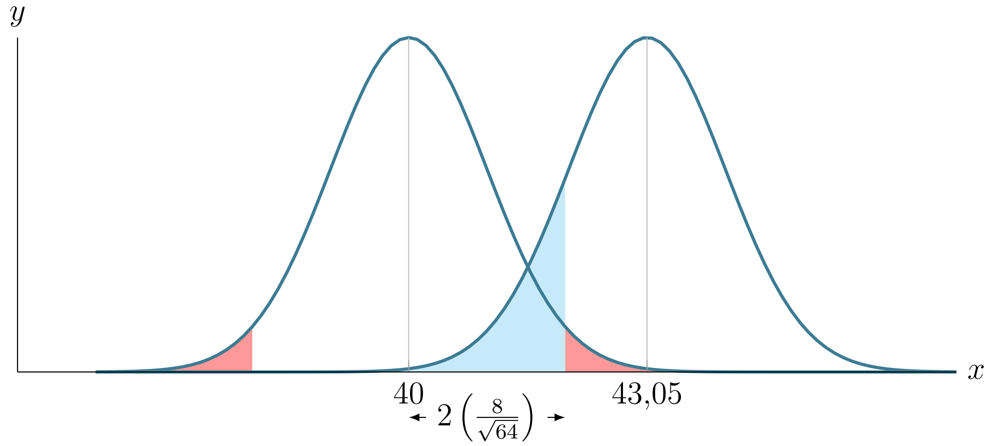
d) (5 puntos) Complete el test de hipótesis de la parte (c) a un nivel de significancia de $\alpha = 0{,}05$.

Solución: Comprobamos la nula $H_0 : \mu = 40$ contra la alternativa $H_1 : \mu \neq 40$. Bajo la nula sabemos que $t = \frac{\hat{\mu}-\mu}{\hat{\sigma}/\sqrt{N}}$ se distribuye según un t con 63 grados de libertad. A un nivel de significancia de $\alpha = 0,05$ entonces rechazamos la nula si $|t| > t_{critico}$, donde $t_{critico} = 2$. Entonces, calculamos que $t = 1{,}67$, y dado que $1{,}67 < 2$, no se rechaza H_0.

e) (2 puntos) Explique la importancia del término $\alpha = 0,05$. ¿Cómo se llama este valor?

Solución: El valor de α define la proporción de veces que estaríamos dispuestos a rechazar la nula cuando la nula es cierta si se podría repetir infinitas veces el experimento subyacente. Se conoce como el error tipo I.

f) (7 puntos) Ahora, imagine que se sabe que el valor verdadero para el parametro poblacional era $\mu = 43,046$. Si se habría comprobado la nula $H_0 : \mu = 40$, cuando en verdad $\mu = 43,046$, ¿cuál es la probabilidad de NO rechazar la nula? [*Pista*: Podría ser útil dibujar las dos distribuciones de interés (la distribución bajo la nula, y la distribución centrada en el valor verdadero) para visualizar cómo contestar a la pregunta].
Nota de corrección: Esta pregunta es difícil. Se puede ganar 6 de 7 puntos sin haber hecho la última parte de encontrar el valor exacto de la área azul. Solución: Si el valor verdadero es 43,046 y comprobamos la nula que $\mu = 40$, no recharacemos la nula siempre cuando $38 < \widehat{\mu} < 42$. Si nos fijamos en el gráfico a continuación, el area azul define la proporción de la curva de densidad del parametro verdadero que caerá abajo del valor 42, y por ende la probabilidad de no rechazar la nula $\mu_{H_0} = 40$ aun cuando $\mu = 43{,}046$. Entonces, la pregunta consiste en saber cuánta de la masa de probabilidad de la curva centrada en 43,046 cae abajo de 42. Si la desviación estándar de ambas curvas es igual a $\hat{\sigma}^2/\sqrt{N} = 1$, sabemos que la distancia entre el promedio de la curva centrada en 43,046 al inicio de la parte azul (42) es igual a $1,046 \times \left(\frac{8}{\sqrt{(64)}}\right)$. Entonces, para encontrar cuánta de la masa de probabilidad cae abajo de este punto, necesitamos encontrar el valor de 1,046 en la tabla t con 60 grados de libertad, que esta en $p = 0{,}85$ Dado que nos interesa el punto de *negativo* 1,046, se concluye que el valor es $1 - 0,85 = 0,15$.



g) (1 puntos) ¿Qué tipo de error se está considerando en el parte (f)?

Solución: Error tipo II (no rechazar la nula cuando la nula es falsa).