Prueba Parcial
Econometría I: Magíster en Economía
Universidad de Santiago de Chile
Semestre 1, 2019

Nombre de alumna/alumno y firma:	
Pregunta 1 (Máximo 12 puntos)	
Pregunta 1 (Máximo 24 puntos)	
Pregunta 2 (Máximo 24 puntos)	

Instrucciones Generales

1. Tiene 90 minutos para responder a esta prueba, dividida en 60 puntos.
2. Póngale nombre a todas las hojas.
3. Se permite el uso de calculadoras, siempre y cuando no cuenten con dispositivos de comunicación.
4. 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. **Repaso de Álgebra Lineal** (12 puntos)

   a) (2 puntos) ¿Qué es una matriz semi-definida positiva (sdp) y definida positiva?

   Solución: Para una matriz $\mathbf{A}$, si $\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$ para cualquier candidato $\mathbf{x}$ (donde $\mathbf{x}$ es un vector), $\mathbf{A}$ se conoce como una matriz semi-definida positiva. Y si $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ para cualquier candidato $\mathbf{x} \neq \mathbf{0}$, $\mathbf{A}$ se conoce como una matriz definida positiva

   b) (4 puntos) Demuestre que la matriz $I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ es definida positiva.

   Solución: Consideramos cualquier vector no cero $\mathbf{x} = [x_1 \; x_2 \; x_3 \; x_4]'$. Y notamos que:

$$Q = \mathbf{x}'\mathbf{I_4}\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_1^2 + x_2^2 + x_3^2 + x_4^2.$$

   Ahora, cada elemento $x_i^2 > 0 \; \forall \; i \in 1, 2, 3, 4 \Rightarrow Q > 0$. Entonces, según la definición de la pregunta 1a, la matriz I_4 es definida positiva.

   c) (3 puntos) Si una matriz de $K \times K$ tiene un inverso, ¿Se puede concluir que no tiene otro inverso? ¿Por qué, o por qué no?

   Solución: Sí, se puede concluir que no tiene otro inverso. Si una matriz es invertible, el inverso es único. Para ver porqué, supongamos que la matriz $\mathbf{A}$ tiene inverso $\mathbf{B}$. Y ahora supongamos que existe otro inverso $\mathbf{C}$. En este caso:

$$(\mathbf{CA})\mathbf{B} = \mathbf{IB} = \mathbf{B} \tag{1}$$

$$\mathbf{C}(\mathbf{AB}) = \mathbf{CI} = \mathbf{C}, \tag{2}$$

   y 1-2 son inconsistentes si $\mathbf{C}$ no es igual a $\mathbf{B}$.

   d) (3 puntos) Si una matriz de $K \times K$ es invertible, ¿Qué se puede afirmar de su rango? ¿Las columnas de esta matriz son linealmente dependientes, linealmente independientes, o no se sabe?

Solución: Se puede afirmar que es de rango completo, y que el rango es igual a K. Por el teorema de inversión matricial, sabemos que las columnas son linealmente independientes.

2. **Pregunta Empírica** [24 puntos] Según las mediciones del *Goddard Institute for Space Studies* (GISS) de NASA, la temperatura promedio en la superficie de la tierra entre 1951-1980 fue 14°C. Ud. cuenta con 20 mediciones de la temperatura promiedo de la tierra entre los años 1991-2010. Suponga que cada valor anual es independiente. El promedio estimado de estos 20 valores es $\hat{\mu} = 14{,}21$°C, y la desviación estándar estimada es $\hat{\sigma} = 0{,}1183$°C.

   a) (8 puntos) Le interesa examinar si hay evidencia estadística que sugiere que la temperatura promedio ha subido en este periodo. Por lo tanto, quiere comprobar la hipótesis nula que el promedio de la temperatura entre 1991-2010 es igual o inferior a 14°C, es decir, $H_0 : \mu \leq 14$. Complete el test de hipótesis a un nivel de significancia de $\alpha = 0{,}05$.

   Solución: Es importante aquí darse cuenta que el test de hipótesis es un test de una sola cola. Comprobamos la nula $H_0 : \mu \leq 14$ contra la alternativa $H_1 : \mu > 14$. Bajo la nula sabemos que $t = \frac{\hat{\mu}-\mu}{\hat{\sigma}/\sqrt{N}}$ se distribuye según una distribución t de Student con 19 grados de libertad. A un nivel de significancia de $\alpha = 0,05$ entonces rechazamos la nula si $t > t_{critico}$, donde el $t_{critico}$ con 19 grados de libertad es: $t_{critico} = 1{,}729$. Aquí es importante notar que rechazamos la nula cuando t es más grande que el valor crítico, pero no cuando t es inferior que $-t_{critico}$, ya que valores bajos de $\hat{\mu}$, (y por ende valores bajos de t), son evidencia a favor de la nula. Entonces, calculamos que $t = (14{,}21 - 14)/(0{,}1183/\sqrt{20}) = 7,94$, y dado que $7{,}94 > 1{,}729$, rechazamos la nula.

   b) (6 puntos) Calcule un estimador del intervalo de confianza de 95 % para el parámetro poblacional en base a los valores observados con las 20 observaciones. Explique en palabras la interpretación de este intervalo de confianza.

Solución: 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 da un valor de 2.093. Entonces, utilizando los valores estimados:

$$\begin{aligned} 14{,}21 \pm 2{,}093 \times \frac{0{,}1183}{\sqrt{20}} \\ 14{,}21 \pm 2{,}093 \times 0{,}0265 \end{aligned} \tag{3}$$

y el estimador de intervalo es [14.155;14.265]. Si se calcula dicho intervalo de confianza infinitas veces en base a muestras representativas de datos, en 95 % de los casos, el parametro poblacional desconocido caerá adentro del intervalo calculado.

c) (5 puntos) En un contraste de hipótesis, existen dos tipos de error (en la decisión de rechazar o no la nula) que se puede cometer. ¿Cómo se definen estos dos errores, y qué serán en el contexto específico de esta pregunta?
Solución: El error tipo I se define como la decisión de rechazar H_0 cuando H_0 es cierta. El error tipo II se define como la decisión de no rechazar H_0 cuando H_0 no es cierta. En este caso específico el error tipo I referería a concluir que la temperatura ha aumentado comparado con el promedio histórico de 14 grados, cuando en realidad no ha aumentado. Y el error tipo II ocurriría si concluimos que la temperatura no ha subido, cuando en realidad, el promedio es más alto.

d) (5 puntos) Imagine que se estima $\hat{\mu}$ dos veces por máxima verosimilitud: uno de forma no restringida, y otro *restringiendo* que $\hat{\mu} = 14$. El valor máximo de la función de log verosimilitud en estos dos casos es (respectivamente) $-19{,}253$ y $-34{,}198$. En este caso ¿Cuál sería el valor del estadístico de prueba en un test de razón de verosimiltudes? ¿Cómo se distribuye esta estadístico de prueba bajo la nula?
Solución: El estadístico de prueba para un test de razón de verosimilitudes es:

$$2[\ell(\hat{\mu}_{MV}) - \ell(\mu_0)] = 2[-19{,}253 + 34{,}198] = 29{,}89.$$

Esta estadística de prueba sigue una distribución χ^2 con 1 grado de libertad *bajo la nula.*

3. **Pregunta Teórica** [24 puntos] Se cuenta con una muestra de tamaño N extraída de una distribución *desconocida*, con media μ y varianza finita σ^2. Las observaciones se denotan como $X_1, X_2, \ldots, X_N$, o en formato vectorial simplemente como $\boldsymbol{X}$. Cada observación es independiente e idénticamente distribuida.

   a) (5 puntos) Sin saber la forma funcional de la distribución subyacente, ¿podemos concluir algo acerca de la distribución límite del promedio muestral $\bar{X}_N$, o de alguna función del promedio muestral? Justifique su respuesta.
   Solución: Sí, por el teorema del límite central de Lindeberg-Lévy sabemos que

$$Z_N = \frac{\bar{X}_N - \mu}{\sigma/\sqrt{N}} \xrightarrow{d} \mathcal{N}(0,1).$$

   Esta variable Z_N es una función de $\bar{X}_N$, especificamente una versión estandarizada.

   b) (7 puntos) En base a los datos, se quiere estimar μ y σ^2. Plantea una metodología para estimar estos parámetros que *no* requiere de supuestos distribucionales, sino simplemente en base a los momentos existentes. Describe el proceso en este caso.
   Solución: Dado el enunciado, solamente sabemos que existen dos momentos poblacionales. Podemos utilizar estos momentos para estimar los parametros de interés utilizando método de momentos. Estos primeros dos momentos son la esperanza y la varianza, los que en la población se escriben:

$$\begin{aligned} E(X) &= \mu & (4) \\ Var(X) &= E[X^2] - (E[X]^2) = \sigma^2. & (5) \end{aligned}$$

   Los mismos momentos centrales son:

$$\begin{aligned} E[X] - \mu &= 0 & (6) \\ E[X^2] - (E[X]^2) - \sigma^2 &= 0. & (7) \end{aligned}$$

y entonces, los vamos a estimar usando las condiciones de momentos centrales análogas de la muestra que observamos:

$$\frac{1}{n}\sum_{i=1}^{n} X_i - \widehat{\mu} = 0 \tag{8}$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2 - \widehat{\sigma}^2 = 0 \tag{9}$$

Son dos momentos y dos parametros a estimar, y por lo tanto es un sistema identificado.

c) (6 puntos) Ahora suponga que adicionalmente, se sabe que $f(X_i|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(\ln(X_i)-\mu)^2}{2\sigma^2}\right)$. Plantee un estimador más eficiente que el estimador del apartado b), y explique cómo estimar los parámetros μ y σ^2 utilizando esta metodología.

Solución: Planteamos un estimador de máxima verosimilitud. Partimos escribiendo la función de densidad de probabilidad conjunta:

$$\begin{aligned} f(X_1,...,X_N|\mu,\sigma) &= f(X_1|\mu,\sigma)\times\cdots\times f(X_N|\mu,\sigma) \\ &= \prod_{i=1}^{N} f(X_i|\mu,\sigma) \\ &= \prod_{i=1}^{N}\frac{1}{\sigma\sqrt{2\pi}}\exp - \frac{(\ln(X_i)-\mu)^2}{2\sigma^2} \end{aligned} \tag{10}$$

donde el último paso viene de la fdp dada en la pregunta. Dado esto, nuestra función de verosimilitud para una muestra de variables extraída de esta distribución es:

$$\mathcal{L}(\mu,\sigma|X_1,\ldots,X_N) = \prod_{i=1}^{N}\frac{1}{\sigma\sqrt{2\pi}}\exp - \frac{(\ln(X_i)-\mu)^2}{2\sigma^2}$$

y el logaritmo de la función de verosimilitud es:

$$\begin{aligned}\ell(\mu,\sigma|X_1,\ldots,X_N) &= N\ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right)-\sum_{i=1}^{N}\frac{(\ln(X_i)-\mu)^2}{2\sigma^2}. \\ &= -\frac{N}{2}\ln(2\pi\sigma^2)-\sum_{i=1}^{N}\frac{(\ln(X_i)-\mu)^2}{2\sigma^2}. \qquad (11)\end{aligned}$$

Esta última ecuación (11) es una función que podríamos entregar a un programa computacional para maximar, y encontrar los parámetros estimados $\hat{\boldsymbol{\theta}}=(\widehat{\mu},\widehat{\sigma})$:

$$\hat{\boldsymbol{\theta}}_{MV}=\arg\max_{\boldsymbol{\theta}}\ell(\mu,\sigma|y_1,\ldots,y_N).$$

d) (6 puntos) Describa brevemente dos propiedades que debemos tomar en cuenta al momento de considerar un estimador. Describa una manera de considerar estas dos propiedades en conjunto al comparar distintos estimadores, y demuestra porqué su metodología toma en cuenta las dos propiedades.

Solución: Se debe considerar la sesgadez/insesgadez de un estimador, y su eficiencia (o la varianza del estimador). Si es un estimador asintótico, debemos considerar su consistencia y eficiencia asintótica. Una medida que considera tanto el sesgo de un estimador y además la precisión del estimador es:

$$ECM(\hat{\theta})=E[\hat{\theta}-\theta]^2 \qquad (12)$$

que incorpora un castigo por sesgo, y además por precisión. Para ver esto, notamos que podemos re-escribir la función de pérdida 12 como:

$$
\begin{aligned}
ECM(\hat{\theta}) &= E[\hat{\theta} - E(\hat{\theta}) + E(\hat{\theta}) - \theta]^2 \\
&= E\{[\hat{\theta} - E(\hat{\theta})] + [E(\hat{\theta}) - \theta]\}^2 \\
&= E[\hat{\theta} - E(\hat{\theta})]^2 + [E(\hat{\theta}) - \theta]^2 + 2E[\hat{\theta} - E(\hat{\theta})][E(\hat{\theta}) - \theta] \\
&= E[\hat{\theta} - E(\hat{\theta})]^2 + [E(\hat{\theta}) - \theta]^2 \\
&= \mathrm{var}(\hat{\theta}) + [\mathrm{sesgo}(\hat{\theta})]^2.
\end{aligned}
$$

Tab. 1: fda Normal Estandarizada

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Tab. 2: Valores Críticos de la Distribución t de Student

k	75.0 %	80.0 %	87.5 %	90.0 %	95.0 %	97.5 %	99.0 %	99.5 %	99.9 %
1	1.000	1.376	2.414	3.078	6.314	12.706	31.821	63.657	318.31
2	0.816	1.061	1.604	1.886	2.920	4.303	6.965	9.925	22.327
3	0.765	0.978	1.423	1.638	2.353	3.182	4.541	5.841	10.215
4	0.741	0.941	1.344	1.533	2.132	2.776	3.747	4.604	7.173
5	0.727	0.920	1.301	1.476	2.015	2.571	3.365	4.032	5.893
6	0.718	0.906	1.273	1.440	1.943	2.447	3.143	3.707	5.208
7	0.711	0.896	1.254	1.415	1.895	2.365	2.998	3.499	4.785
8	0.706	0.889	1.240	1.397	1.860	2.306	2.896	3.355	4.501
9	0.703	0.883	1.230	1.383	1.833	2.262	2.821	3.250	4.297
10	0.700	0.879	1.221	1.372	1.812	2.228	2.764	3.169	4.144
11	0.697	0.876	1.214	1.363	1.796	2.201	2.718	3.106	4.025
12	0.695	0.873	1.209	1.356	1.782	2.179	2.681	3.055	3.930
13	0.694	0.870	1.204	1.350	1.771	2.160	2.650	3.012	3.852
14	0.692	0.868	1.200	1.345	1.761	2.145	2.624	2.977	3.787
15	0.691	0.866	1.197	1.341	1.753	2.131	2.602	2.947	3.733
16	0.690	0.865	1.194	1.337	1.746	2.120	2.583	2.921	3.686
17	0.689	0.863	1.191	1.333	1.740	2.110	2.567	2.898	3.646
18	0.688	0.862	1.189	1.330	1.734	2.101	2.552	2.878	3.610
19	0.688	0.861	1.187	1.328	1.729	2.093	2.539	2.861	3.579
20	0.687	0.860	1.185	1.325	1.725	2.086	2.528	2.845	3.552
21	0.686	0.859	1.183	1.323	1.721	2.080	2.518	2.831	3.527
22	0.686	0.858	1.182	1.321	1.717	2.074	2.508	2.819	3.505
23	0.685	0.858	1.180	1.319	1.714	2.069	2.500	2.807	3.485
24	0.685	0.857	1.179	1.318	1.711	2.064	2.492	2.797	3.467
25	0.684	0.856	1.178	1.316	1.708	2.060	2.485	2.787	3.450
26	0.684	0.856	1.177	1.315	1.706	2.056	2.479	2.779	3.435
27	0.684	0.855	1.176	1.314	1.703	2.052	2.473	2.771	3.421
28	0.683	0.855	1.175	1.313	1.701	2.048	2.467	2.763	3.408
29	0.683	0.854	1.174	1.311	1.699	2.045	2.462	2.756	3.396
30	0.683	0.854	1.173	1.310	1.697	2.042	2.457	2.750	3.385
35	0.682	0.852	1.170	1.306	1.690	2.030	2.438	2.724	3.340
40	0.681	0.851	1.167	1.303	1.684	2.021	2.423	2.704	3.307
45	0.680	0.850	1.165	1.301	1.679	2.014	2.412	2.690	3.281
50	0.679	0.849	1.164	1.299	1.676	2.009	2.403	2.678	3.261
55	0.679	0.848	1.163	1.297	1.673	2.004	2.396	2.668	3.245
60	0.679	0.848	1.162	1.296	1.671	2.000	2.390	2.660	3.232
100	0.677	0.845	1.157	1.290	1.660	1.984	2.364	2.626	3.174
∞	0.674	0.842	1.150	1.282	1.645	1.960	2.326	2.576	3.090