Prueba Parcial – PAUTA
Econometría I: Magíster en Economía
Universidad de Santiago de Chile

Nombre de alumna/alumno y firma:	
Pregunta 1 (Máximo 30 puntos)	
Pregunta 2 (Máximo 30 puntos)	
Pregunta 3 (Máximo 30 puntos)	

Instrucciones Generales

1. Tiene 90 minutos para responder a esta prueba, dividida en 90 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.

Buena suerte!!

1. **Varias Preguntas** [20 puntos]

   a) (5 puntos) El error cuádratico medio (ECM) es una estadística que combina la precisión de un estimador con su sesgo. El ECM se define como: $ECM = E[\hat{\beta} - \beta]^2$, donde β es una cantidad poblacional, y $\hat{\beta}$ su estimador. Demuestra como el ECM depende de los dos aspectos de un estimador: la precisión y el sesgo.
   Es necesario mostrar que el ECM puede ser expresado como una función de la varianza de un estimador, y también su sesgo (si solo se explica esto, 2 puntos). Vimos en la clase de estimadores que podemos hacer la siguente re-organización:

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

   y de la última línea, vemos que el ECM depende de la varianza, y el sego del estimador al cuadrado.

   b) (10 puntos) Una institución engargada de analizar el sangre de los/as deportistas para ver si hay algún tipo de irregularidad le contrata para analizar la veracidad estadística de su tecnología. Le informa que han desarrollado un nuevo tipo de test analítico que logra identificar la presencia de drogas ilegales en el sangre en 99 % de los casos. Pero a la vez, en 0,01 % de los casos, el test indica que hay una presencia de drogas cuando en realidad no lo hay (es decir, una positiva falsa). Imagina que sabemos que 1 de cada 1,000 deportistas está consumiendo drogas ilegales. Si una deportista fue eligido aleatoriamente de la población y el test arroja un resultado positivo, ¿cuál es la probabilidad de que la persona efectivamente está consumiendo drogas ilegales? Asegure de mostrar los cálculos, y nombre brevemente cualquier teorema o fórmula que aplique para llegar a la solución.
   Para resolver esta pregunta es necesario utilizar el teorema de Bayes (2 puntos). Este teorema nos da que: $Pr(B_i|A) = \frac{Pr(B_i)Pr(A|B_i)}{\sum_{j=1}^{k} Pr(B_j)Pr(A|B_j)}$ (2 puntos). Aquí A es que la persona dio un test positivo, y B_i

es que la persona efectivamente consumio drogas. Sabemos cual es la probabilidad de B_i (1/1000), y la probabilidad de haber arrojado un test positivo dado que la persona consumió drogas ($A|B_i$=0.99). Podemos calcular el $Pr(B_i|A)$ como:
0.001×0.99/(0.001×0.99+0.999×0.001)=0.908.

c) (5 puntos) Una empresa de productos electrónicos asegura que sus productos tienen una vida útil promedio de 2 años (730 días), y una varianza de 100 días. Sin embargo, después de comprar un producto electrónico de esta empresa, el producto falla después de 30 días. ¿Qué se puede decir acerca de la probabilidad de este evento?
Si se identifica que este problema se puede resolver mediante la desigualdad de Chebyshev, 1 punto. Si se identifica la desigualdad de Chebyshev como $Pr(|X - E(X)| \geq t) \leq Var(X)/t^2$., 1 punto. La respuesta entonces es: $Pr(|X - E(X)| \geq 700) = 100/700^2 =$ 0,0002. (2 puntos). Se puede decir que la probabilidad de haber observado este evento es muy bajo.

2. **Estimadores** [35 puntos] Para esta pregunta, vamos a considerar una variable aleatoria X, y asumimos que $X \sim \mathcal{N}(\mu, \sigma^2)$. La distribución normal tiene una función de distribución de probabilidad: $f(x|\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$. Imagine que tiene una muestra iid de N realizaciones de X, cada uno denominado $x_i \; \forall \; i \in 1, \ldots, N$.

   a) (12 puntos) Deriva una estimador de maxima verosimilitud que le permite estimar los dos parámetros de la distribución normal (μ y σ^2). Es necesario escribir la función a mazimizar, pero *no* es necesario buscar una solución analítica al problema de maximización. Una respuesta completa partiría con la fdp de una normal $f(x_i|\mu,\sigma^2)$, y llegaría a la fdp conjunta multiplicando la fdp de todas las N fdp individuals, dado que cada las realizaciones son iid. Después es necesario indicar que la función de verosimilitud se escribe como $\mathcal{L}(\mu,\sigma^2|x_1,\ldots,x_N)$, y que tomamos el logaritmo de esto para formar la función de log verosimilitud. Estimamos los parametros ($\hat{\mu}$ y $\hat{\sigma}^2$) maximizando esta función.

   b) (3 puntos) En la función que derivó en la parte *a*, cuáles son los argumentos de la función sobre cuál estamos maximizando? ¿Cuál es la función objetiva?
   Los argumentos sobre que estamos maximizando son los parametros de la normal. La función objetiva es la función de log verosimilitud. También se acepta la función de verosimilitud como respuesta a la función objetiva.

   c) (2 puntos) Y cuál es el rango de esta función objetiva?
   Si se dice la función de log verosimilitud en la parte anterior, entonces: $(-\infty, 0)$. Si se indica que estamos utilizando la función de verosimilitud, entonces: (0,1).

   d) (5 puntos) Explique brevemente la diferencia entre estimadores de maxima verosimilitud (MV) y estimadores de método de momentos (MM). La diferencia está en cómo se motiva estimación. En MV se elige los valores maximizando toda la fdp conjunta, mientras en MM solo se trata de encontrar los valores de los parametros para satisfacer algunos momentos de la distribución.

   e) (8 puntos) Ahora, imagine que en vez de estimar por MV, ud quiere estimar los dos parámetros de interés (μ y σ^2) utilizando MM.

¿Cuáles son los momentos poblacionales que sabemos que cumplen? ¿Y cuáles son las contrapartes muestrals que se ocupa para estimar los parametros? Para un estimador MV, los dos momentos poblacionales más fáciles de utilizar son la media y la varianza:

$$E(Y) - \mu = 0$$

$$Var(Y) = E[Y^2] - (E[Y]^2) - \sigma^2 = 0.$$

Para poder estimar estos momentos, utilizamos sus dos contrapartes muestrales:

$$\frac{1}{n}\sum_{i=1}^{n} Y_i - \hat{\mu} = 0$$

$$\frac{1}{n}\sum_{i=1}^{n} Y_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right)^2 - \hat{\sigma}^2 = 0$$

Donde eligimos $\hat{\mu}$ y $\hat{\sigma}^2$ para hacer cumplir estos dos momentos muestrales.

f) (3 puntos) ¿Cuál de estos dos estimadores (MV o MM) tiene supuestos más exigentes? ¿Por qué? MV. En este caso necesitamos hacer un supuesto acerca de toda la distribución de la variable aleatoria en la población, mientras con MM, solo estamos haciendo supuestos acerca de un par de momentos.

g) (2 puntos) ¿Utilizando MM, sería posible también estimar la kurtosis de nuestra variable aleatoria? ¿Por qué sí, o por qué no? En la teoría, sí, podemos estimar cualquier valor poblacional si es que contamos con por lo menos tantos condiciones de momentos como parametros para estimar. En la práctica, probablemente no es una cantidad muy interesante, ya que creemos que es una variable normal, y por lo tanto, kurtosis=0.

3. **Aplicación Empírica** [35 puntos] El Indíce de Calidad del Aire referido a Partículas (ICAP) es una medida de la contaminación en el aire y el daño que ésta puede provocar a las personas. Un valor sobre 300 implica que hayan fuertes riesgos a la salud. Imagine que ud cuenta con una muestra (de $N = 30$) de datos recolectado en todo el año 2015 en Santiago que mide el ICAP en determinadas días. El promedio en tu muestra $\hat{\mu}_{ICAP}$=286,5, y la desviación estándar $\hat{\sigma}_{ICAP}$=40,72. Examinando la distribución de tu variable, ud tiene motivo de creer que la variable subyacente está distribuída según una normal.

   a) (6 puntos) Calcule un estimador del intervalo de confianza de 95 % para el ICAP promedio en Santiago en base de los datos que tiene a mano. Explicita cualquier supuesto que tiene que hacer para construir este estimador. Para construir un intervalo de confianza de 95 %, podemos utilizar la distribución t si hacemos el supuesto de que cada muestra es independiente e idénticamente distribuida. Para calcular el intervalo de confianza utilizamos la media, la cantidad de observaciones, y la desviación estándar del enunciado. Para calcular el percentil de la distribución, tenemos que mirar en la tabla t en el punto con 28 grados de libertad (se pierde dos por la media y el desviación estándar). El intervalo final es: $286,5 \pm 2{,}045 \times 40,72/\sqrt{30} = 286,5 \pm 15{,}2$

   b) (12 puntos) Testee formalmente si $\mu_{ICAP} < 300$ a un nivel de significancia de 99 %. Asegure formular la hipótesis nula, la alternativa, el estadístico de prueba, la región de rechazo, y el resultado del test. Nuestra test es: $H_0: \ \mu_{ICAP} < 300$ vs. $H_1: \ \mu_{ICAP} \geq 300$. Utilizando la tabla t, tenemos 28 grados de libertad, y estamos intersado en un test de una cola. Nos da un t-critico de 2,462. Ahora, utilizando los valores observados, podemos calcular nuestra estadistica de prueba como: $(286,5 - 300)/(40,72/\sqrt{30}) = -1{,}81$. Dado que el valor absoluto de la estadística de prueba es menor que el valor crítico $(1,81 < 2,462)$, no se rechaza la nula.

   c) (3 puntos) ¿Cuál es el valor p asociado con este test de hipótesis (Explique el proceso de calcularlo, y si es necesario, redondee tu respuesta al valor más cercano posible)? Para calcular el valor p, tenemos que identificar la masa de la distribución de probabilidad que cae arriba de 1,81 en una distribución t con 28 grados de libertad. De la tabla estadística, vemos que el percentil 95 está en

1,701, y el percential 97,5 está en 2,048. Entonces, sabemos que el valor p es un poco más alto que el percentil 95, o sea, un valor un poco más bajo que 1-0,95=0,05.

d) (3 puntos) En base a las respuestas anteriores, explique cuál es el error tipo I en este caso. ¿Qué relación tiene este error con tu respuesta en la parte b de esta pregunta? El error tipo I consiste en rechazar la nula cuando la nula es cierta. En este caso, sería rechazar que el nivel del ICAP promedio es menor de 300 cuando en realidad, no lo es. La probabilidad de hacer un error tipo I es igual al nivel de significancia del test. En la parte b, hay un 1 % de probabilidad de hacer un error tipo I.

e) (3 puntos) En términos generales, ¿Cuáles son los costos asociados con un error tipo I y un error tipo II? ¿En este caso, cómo debemos formar nuestro nivel de significancia cuando pensamos en los costos reales de cada tipo de error? El costo de cada tipo de error depende del contexto. Un error tipo I es rechazar la nula cuando la nula está verdadera, y un error tipo II consiste en no rechazar la nula cuando la nula es falsa. Si pensamos que es muy costoso hacer un error tipo I, debemos definir un nivel de significancia muy exigente.

f) (8 puntos) Ahora imagine que ud cuenta con una muestra para el mismo periodo (y el mismo tamaño de muestra) de la ciudad de Concepción, donde $\hat{\mu}^c_{ICAP}$=223,4 y $\hat{\sigma}^c_{ICAP}$=24,80. Proponga una manera para testear si existe un diferencia estadísticamente significativo entre las dos ciudades (con un sólo test de hipótesis). Realiza el test de hipótesis propuesta a un nivel de significancia de 95 %, siguiendo los mismos pasos descritos en la parte b de esta pregunta. *Pista: Si se quiere testear una hipótesis para comparar dos medias, se puede calcular el denominador del test de hipóteis como:* $\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$. *Además, los grados de libertad en este caso están aproximados por 28.* Esta es una pregunta avanzada, ya que hasta ahora no hemos visto test de hipótesis de este tipo. Ahora queremos testear si $\mu_{ICAP} = \mu^c_{ICAP}$. Esto implica las hipótesis: $H_0: \ \mu_{ICAP} = \mu^c_{ICAP}$ vs. $H_1: \ \mu_{ICAP} \neq \mu^c_{ICAP}$ Para hacerlo más simple, se puede transformar la nula en $H_0: \ \mu_{ICAP} - \mu^c_{ICAP} = 0$. Una manera de construir una estadística de prueba entonces es

$t = \frac{\mu_{ICAP} - \mu^c_{ICAP} - 0}{\sqrt{\sigma_c^2/n_c + \sigma^2/n}}$, donde c refiere a los valores para Concepción, y los otros valores son de Santiago. Entonces, para calcular el valor de la estadística de pruebatenemos: $\frac{286{,}5 - 223{,}4 - 0}{\sqrt{24{,}80^2/30 + 40{,}72^2/30}} = 7{,}248$. De la tabla t, tenemos un valor t crítico de 95 % (2 colas) y 28 grados de libertad de 2,048. Entonces, como $7,248 > 2,048$, se rechaza la nula. NOTE: para ganar todos los puntos de esta pregunta no es necesario tener todos los pasos y álgebra correctos, dado que la naturaleza más avanzada de la pregunta.