

[Página Principal](#) / [Mis cursos](#) / [Carreras de Grado](#) / [Materias Comunes](#) / [Período Lectivo 2023](#) / [Estadística 2023](#)
 / [EVALUACIONES 2023 | 01](#) / [SEGUNDO PARCIAL - Comisión N° 1 - Mañana - 21/06/2023 10 hs.](#)

Comenzado el	Wednesday, 21 de June de 2023, 10:30
Estado	Finalizado
Finalizado en	Wednesday, 21 de June de 2023, 12:29
Tiempo empleado	1 hora 59 minutos
Puntos	12,00/20,00
Calificación	6,00 de 10,00 (60%)

Pregunta 1

Finalizado

Se puntúa 1,70 sobre 3,33

Para fabricar cierto microchip se compararon dos procesos. De un proceso menos costoso se seleccionaron 400 chips, de los cuales se sabía que 62 tenían algún defecto. De un proceso más costoso, se seleccionaron 100 chips, de ellos 12 tenían defecto.

Se están considerando tres planes de muestreo de igual costo. En el primer plan se probarán 100 chips más del proceso menos costoso. En el segundo se observaron 50 más del proceso más costoso. En el tercer plan se probaron 50 del menos costoso y 25 del más costoso. ¿Qué plan es más probable que proporcione el aumento más grande en la precisión del intervalo de confianza? Explique.

P1:proceso menos costoso

P2:proceso mas costoso

Plan 1: 100 chips + de P1. = 500 chips

Proporción de chips defectuosos = $62/400$, no se agregarían más chips defectuosos.

Plan 2: 50 chips + de P2. = 150 chips

Proporción de chips defectuosos $12/100$, no se agregarían más chips defectuosos.

Plan 3: 50 de P1 y 25 de P2.

475 chips de P1 y 125 chips de P2.

Proporción de chips defectuosos $62/400$ para P1 y $12/100$ para P2

El aumento en la precisión del intervalo de confianza está relacionado con la raíz cuadrada del tamaño de la muestra. Cuanto mayor sea el tamaño de la muestra, menor será la amplitud del intervalo de confianza y, por lo tanto, mayor será la precisión.

Comparando los tres planes, el plan 1 es el que mayor tamaño de muestra nos brinda (500), entonces el plan 1 sería el mejor para proporcionar el aumento en la precisión del intervalo de confianza.

Igualmente el intervalo de confianza también estaría influenciada por la proporción de chips defectuosos pero como no se agregan casos defectuosos el plan se basa en el tamaño de muestra.

Comentario:

Está bien la idea y vale para, por ejemplo, un IC para una media. Pero en este caso el IC es más complejo y un tamaño de muestra total mayor no se traduce directamente en un aumento de la precisión, ya que depende (como menciona) de las proporciones p_1 y p_2 , aunque están fijas. Tendría que haber calculado los errores de cada IC y compararlos entre sí.

Revisión:

Es importante considerar que en un ejercicio de práctica, la precisión puede calcularse y no responderse en forma abstracta.

Revisado por el Prof. Juan.

Pregunta 2

Finalizado

Se puntúa 2,50 sobre 3,33

Los óxidos de nitrógeno (Nox) constituyen un componente muy importante de la contaminación atmosférica producida por vehículos automotores. En un estudio para determinar la relación entre la carga sobre un motor y la producción de Nox, un vehículo fue manipulado en un laboratorio de prueba a diferentes velocidades. Se realizaron mediciones periódicas de los caballos de fuerza [x] y de las emisiones de Nox [y]. Las emisiones de Nox se midieron en mg/s. El siguiente resultado de un software de estadística describe el ajuste de un modelo lineal con estos datos. Suponga que los supuestos son válidos.

The regression equation is
 Chassis NOx = 44.5 + 0.845 Chassis HP

Predictor	Coef	SE Coef
Constant	44.534	2.704
Chassis HP	0.84451	0.03267

S = 24.62 R-Sq = 84.7% R-Sq(adj) = 84.5%

A) Pronostique o prediga la tasa de emisión de Nox si el motor se detiene a los 10 caballos de fuerza.

B) Explique qué significan los valores en la columna de SE Coef

a)

La ecuación de regresión es

CNox= Chassis NOx

CHP= Chassis HP

CNox= 44.5 + 0.845 * CHP

Si el motor se detiene a los 10 caballos de fuerza debemos tomar como CHP=10

Reemplazando en la formula:

CHP=44.5 +(0.845 * 10)

CHP=52,95

Por tanto el pronostico de la tasa de emisión de NOX si el motor se detiene a los 10 caballos de fuerza es 52,95.

b) Los valores de la columna SE Coef se refieren al error estandar donde

la constante de la ecuacion de regresion tiene 2.704

y Chassis HP tiene 0.03267

donde el error estandar es una medida de la precisión o variabilidad de una estimación estadística. Es decir indica cuanto puede variar la estimación de la media de una población si se toman múltiples muestras de esa población. Como sería en este caso si tomamos múltiples mediciones periódicas de los caballos de fuerza y de las emisiones de Nox.

Se calcula dividiendo la desviación estándar de la población por la raíz cuadrada del tamaño de la muestra. Cuanto mayor sea el tamaño de la muestra, menor será el error estándar y como será menor el error estándar, mayor será la precisión

Comentario:

Bien en general, pero realiza un pronóstico puntual.

Revisado por el Prof. Juan.

Pregunta **3**

Finalizado

Se puntúa 1,80 sobre 3,34

Se dispone de los datos de las ventas en millones de una empresa de IT. Realizar una gráfica de las componentes de la serie separadas.

Trimestres \ Años	2006	2007	2008	2009	2010
Primero	2	3	2	4	5
Segundo	2	4	4	5	6
Tercero	3	5	5	7	8
Cuarto	3	4	4	3	5

..

 [.parcial2_estadistica_bargas.gnumeric](#)

Comentario:

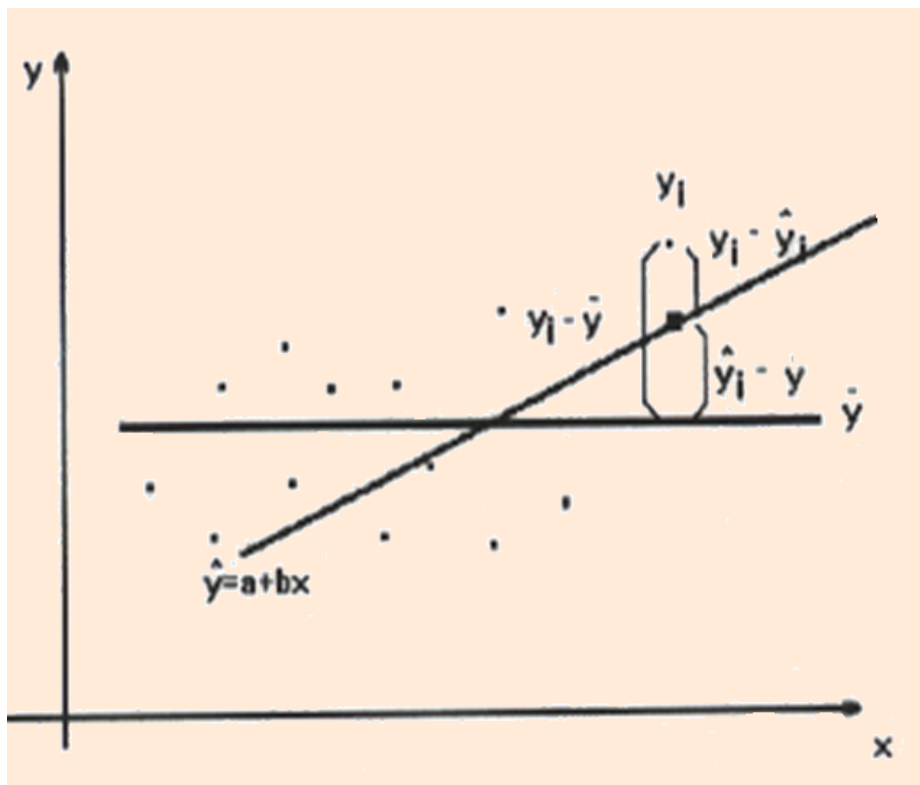
Bien el trabajo con la estacionalidad. No avanza con el resto de las componentes. En cuanto a las gráficas, y las tendencias por año y por cuatrimestre, no es algo que hayamos visto en clase. Te recomiendo que lo veamos en consultas de cara al futuro de tu cursado.

Revisado por el Prof. Juan.

Pregunta 4

Finalizado

Se puntúa 5,00 sobre 5,00



Desarrolle brevemente la teoría que la imagen precedente le sugiere.

Para comenzar en esta grafica vemos la primer recta Ymedia, que es el valor medio

Luego vemos la línea recta representada por $\hat{y} = a + bx$ que es la ecuación de regresión lineal, es decir la recta estimada. Donde la ecuación de regresión lineal es una herramienta estadística utilizada para modelar la relación entre una variable dependiente y una o más variables independientes.

Podemos ver como las observaciones (puntos negros en el grafico) se dispersan en el plano xy

En la ecuación de regresión lineal::

y representa la variable dependiente que queremos predecir.

x es la variable independiente o predictora.

b: los coeficientes de regresión que representan el cambio esperado en y cuando x cambia de unidad

a : es el término de intersección o el valor esperado de y cuando la/las variables independientes son cero.

El objetivo de la regresión lineal es encontrar los mejores coeficientes de regresión que minimicen la suma de los errores cuadrados entre las observaciones reales y las predicciones realizadas por el modelo.

También podemos hablar de r^2 que es el coeficiente de regresión lineal donde

$r^2 = \text{Suma de cuadrados explicados} / \text{suma de cuadrados total}$

Suma de cuadrados explicados = SCE

Suma de cuadrados total = SCT = $\sum (Y_i - \bar{Y})^2$

Sabemos que r^2 varía entre 0 y 1, ya que SCE es menor o igual que SCT.

Podemos decir que:

$SCE = 0 \rightarrow SCR = SCT \rightarrow r^2 = 1$. Esto significa que todos los puntos están sobre la recta estimada.

Si $SCR = 0 \rightarrow SCE = SCT \rightarrow r^2 = 0$. Esto significa que la pendiente de la recta es igual a cero. Esto puede deberse a que la línea de regresión sea horizontal. Y también esto podría ser debido a que las observaciones se dispersan alrededor del valor medio en forma aleatoria o que las observaciones se dispersan alrededor de una curva donde la línea mejor ajustada es una línea recta horizontal, o también, todas las observaciones tienen el mismo valor, cualquiera sea el valor de x

Comentario:
MUY BIEN!!!!

Pregunta **5**

Finalizado

Se puntúa 1,00 sobre 5,00

¿Cuál sería el motivo de asegurar que las muestras grandes son siempre una buena idea?

Decir que las muestras grandes son SIEMPRE buena idea no sería lo apropiado, las muestras grandes pueden ser beneficiosas en muchos casos, pero no siempre en todas las ocasiones. A continuación daré razones porque son buena idea y porque no lo es siempre.

Buena idea:

Al tomar una muestra grande obtenemos una mayor representatividad, es decir los datos que buscamos representar serán mas representativos de la población que se estudia, cuanto mayor sea la muestra mas se reducirá el sesgo() y la posibilidad de obtener resultados distorsionados. Además, la muestra grande proporciona una mayor precisión en los análisis ya que las estimaciones e intervalos de confianza serán mas estrechos. También podemos encontrar relaciones o efectos reales, o diferencias pequeñas entre los grupos.

No siempre buena idea:

Una muestra grande puede ser costoso en recursos como financieros o de tiempo, si la muestra es demasiado grande puede requerir mas recursos de los que se disponen. Como también puede llevar mucho tiempo recopilar una muestra grande, retrasando así los resultados y análisis de los datos.

Podría también haber un error de muestreo donde el error de muestreo se refiere a la diferencia entre los resultados obtenidos de una muestra y los resultados que se obtendrían si se hubiera encuestado a toda la población. También una gran muestra aumenta la dificultad en el manejo y análisis de datos.

Entonces la decisión sobre el tamaño de la muestra debe basarse en una evaluación de los objetivos, recursos disponibles y que tan aptos estamos para realizar muestras.

Comentario:

cuanto mayor sea la muestra mas se reducirá el sesgo() y la posibilidad de obtener resultados distorsionados. Además, la muestra grande proporciona una mayor precisión en los análisis ya que las estimaciones e intervalos de confianza serán mas estrechos. NO SE ENTIENDE LO QUE QUIERE DECIR

También podemos encontrar relaciones o efectos reales, o diferencias pequeñas entre los grupos. Y ACA?????

BUENO NO LOGRÉ CON TODO LO QUE ESCRIBE DECIR EL POR QUÉ. ESTUDIE EL CONCEPTO.

◀ [PRIMER PARCIAL - Comisión N° 1 - Mañana - 10/05/2023 10 hs.](#)

Ir a...

[Tablas de Inferencia ▶](#)