

## Estadística Computacional Certamen I. Lu. 20.10.08

1. (40 puntos) Las siguientes tablas muestran información relativa a la prueba SIMCE 2007 (fuente: Ministerio de Educación).

PUNTAJES PROMEDIO 4º Y 8º BÁSICO POR GRUPO SOCIOECONÓMICO Y EVALUACIÓN DOCENTE						
GRUPO SOCIOECONÓMICO	4º BÁSICO			8º BÁSICO		
	0	1-2	3-4	0	1-2	3-4
BAJO	218	(+)224	(++)230	229	(+)236	(++)238
MEDIO BAJO	218	(+)224	(++)233	231	(+)237	(++)242
MEDIO	233	(+)239	(++)245	241	(+)248	(++)254
MEDIO ALTO	261	(+)273	(+)269	—	282	285
ALTO	—	—	—	—	—	—
TOTAL	223	230	238	234	241	249

0 : Ningún profesor fue bien evaluado.  
1-2 : 1 a 2 profesores fueron bien evaluados.  
3-4 : 3 a 4 profesores fueron bien evaluados.

Figura 1 Puntajes de 4º y 8º básico por grupo socio-económico (5 estratos) y según resultados de la evaluación docente (3 estratos), en colegios municipales.

GRUPO SOCIOECONÓMICO	CURSOS DE 4º BÁSICO						CURSOS DE 8º BÁSICO					
	ESTUDIANTES EVALUADOS						ESTUDIANTES EVALUADOS					
	NÚMERO DE PROFESORES BIEN EVALUADOS						NÚMERO DE PROFESORES BIEN EVALUADOS					
	0		1-2		3-4		0		1-2		3-4	
	N	%	N	%	N	%	N	%	N	%	N	%
BAJO	4.662	4%	6.915	7%	5.531	5%	3.852	3%	12.083	11%	2.589	2%
MEDIO BAJO	15.793	15%	21.646	21%	16.851	16%	13.431	12%	37.120	33%	7.413	7%
MEDIO	6.600	6%	12.379	12%	12.008	11%	6.154	5%	20.477	18%	7.262	6%
MEDIO ALTO	572	1%	981	1%	1.646	2%	472	0%	1.257	1%	1.438	1%
ALTO	—	—	—	—	—	—	—	—	—	—	—	—
TOTAL	27.627	26%	41.921	40%	36.036	34%	23.909	21%	70.937	62%	18.702	16%

0 : Ningún profesor fue bien evaluado.  
1-2 : 1 a 2 profesores fueron bien evaluados.  
3-4 : 3 a 4 profesores fueron bien evaluados.  
— : Indica que la categoría no tiene estudiantes

NOTA : Dado que los porcentajes están aproximados, la suma de los porcentajes de cada categoría puede no coincidir con los porcentajes totales.

Figura 2 Número de alumnos evaluados por grupo socio-económico (5 estratos) y según resultados de la evaluación docente (3 estratos), en colegios municipales.

- (a) Construya estadísticas de tendencia y dispersión para el puntaje SIMCE por estrato económico sin considerar la evaluación segmentación por evaluación docente.
- (b) Construya Box-Plots que permitan estudiar la relación entre evaluación docente y puntajes obtenidos en la prueba SIMCE. Analice y Concluya.
- (c) ¿Qué porcentaje de la varianza en los puntajes obtenidos por los alumnos de 8° básico puede ser atribuida a la evaluación de sus profesores? Concluya. Hint: puede calcular el puntaje medio y varianza en cada una de las 3 clases de evaluación docente.

**2. (35 puntos)** Chile, Marzo 2010. El nuevo gobierno ha lanzado un programa de salud pública consistente en la instalación de dispensadores de condones en los baños públicos de parques, bares, restaurantes y estaciones de servicio. Para permitir su distribución a bajo costo (100 pesos la unidad) se han inaugurado una marca estatal (jaguars ®) con 3 plantas  $A, B, C$  responsables de producir el 45 %, 30 % y 25 % del total respectivamente. Se ha calculado que la probabilidad de que las plantas entreguen un condón con fallas es del 5 %, 15 % y 20 % respectivamente.

- (a) Si compramos 10 unidades, ¿cuál es la probabilidad de que al menos uno salga malogrado?. ¿En este caso, de qué planta es más probable que venga?.
- (b) Marzo 2011. Para mejorar la calidad, los condones ahora se empacan y someten a pruebas de manera centralizada, mediante un procedimiento que ha mostrado fallas inferiores al 10 %. ¿Con qué probabilidad un condón no es rechazado por la planta empacadora?
- (c) Después de implantado este último procedimiento, ¿Cuál es la probabilidad de el ilustre ciudadano obtenga un condón con fallas en la máquina dispensadora?

**3. (25 puntos)** Muchos estudios buscan relacionar variables del entorno con enfermedades y así dar con las causas de éstas y así poder tomar un plan de acción adecuado. En el siguiente caso, se tomaron muestras de agua y se midió su grado de contaminación. Además se guarda un registro del promedio de personas enfermas que pertenecen al área de donde se han tomado las muestras de agua. Los datos son los siguientes:

Zona	A	B	C	D	E	F	G	H	I	J	K	M
Nivel de Contaminación	1	8	3	15	18	5	9	14	2	21	4	18
Promedio enfermos	2	185	20	835	1378	63	240	699	425	1983	37	1539

- (a) Proponga un modelo de regresión lineal que explique la incidencia de la enfermedad en términos de la contaminación del agua. Si es necesario, aplique una transformación a los datos.
- (b) Evalúe si el modelo obtenido modela de manera razonable los datos.
- (c) ¿Aprecia algún dato atípico que podría influenciar la conclusión anterior? Explique.
- (d) Si el muestreo de las aguas se hizo por conveniencia, ¿qué validez tienen las conclusiones obtenidas?