

17.871

Primavera 2002

Boletín de problemas 2: estadística descriptiva y regresiones bivariadas

Repartido: 21 de febrero de 2002

Entregar: 28 de febrero de 2002

#### Parte A.

Resuelva los ejercicios de repaso del texto de Freedman, et al., 3ª edición que se indican más abajo. Cuando un ejercicio requiera una explicación, hágalo de manera legible, usando frases completas, párrafos, etc. Redacte la explicación siempre que ocupe más de 3 líneas.

Capítulo 3: ejercicios de repaso 1, 2, 4, 8, 9, 10 (págs. 50–55)

Capítulo 4: ejercicios de repaso 1, 2, 3, 5, 6, 7, 10 (págs. 74–76)

Capítulo 8: ejercicios de repaso 1, 2, 3, 6, 7, 11 (págs. 134–139)

#### Parte B.

Los datos que aparecen en la página 3 son un muestreo al azar tomado de respuestas de la guía de evaluación de la asignatura de la primavera de 1995.

1. ¿Cuál es el coeficiente de correlación entre la calificación global de la asignatura y el número de estudiantes matriculados en ella?
2. ¿Cuál es el coeficiente de correlación entre la calificación global de la asignatura y el registro del número de estudiantes matriculados en ella?
3. Describa qué correlación resume mejor la relación entre la cantidad de estudiantes matriculados en la asignatura y la calidad de la clase.

Para esta parte, muestre los cálculos obtenidos de sus resultados para B-1 y B-2.

#### Parte C.

Cree un conjunto de datos STATA® de los datos que ha analizado manualmente en la parte B.

1. ¿Cuál es el cálculo de STATA® de la correlación entre la calificación global de la asignatura y el número de matrículas (con 3 cifras decimales)?
2. ¿Cuál es el cálculo de STATA® de la correlación entre la calificación global de la asignatura y el registro (**log**) de estudiantes matriculados en ella (con 3 cifras decimales)?
3. Cree dos gráficas que muestren:
  - a. La calificación global de la asignatura comparada con el número de matrículas.
  - b. La calificación global de la asignatura comparada con el registro (**log**) del número de matrículas.

¿Le obliga esta gráfica a cambiar la respuesta dada en B-3? ¿Por qué o por qué no? (Pista 1: cuando se pide "crear una gráfica comparativa de y

con  $x$ ";  $y$  es la variable dependiente,  $x$  es la variable independiente. Pista 2: al dibujar la gráfica b. emplee la opción ".xlog" en el comando de la gráfica en vez de generar el registro del número de matrículas y trazarlo sobre él.)

Para esta parte, entregue un archivo log que contenga (1) los nombres de la variable de los conjuntos de datos, (2) las medias de las variables, y (3) las correlaciones entre las variables relevantes. Entregue también una versión impresa de la gráfica C-3.

Parte D.

En el presente curso existen dos conjuntos de datos. Ejemplos:

fla\_precinct\_subset.dta: contiene datos electorales básicos, en el ámbito de las circunscripciones, de las elecciones presidenciales de 2000 en Florida.

fla\_county\_subject.dta: contiene datos básicos sobre las técnicas de recuento de votos empleadas en las elecciones presidenciales del 2000 en cada condado.

Existe una descripción de las variables de cada conjunto de datos en la página 4.

Usando estos dos conjuntos de datos, indique las respuestas a las siguientes preguntas:

1. ¿Qué coeficiente de correlación hubo entre el porcentaje de votantes de raza negra registrados en una circunscripción y el porcentaje de papeletas, en cada circunscripción, con votos a más de un candidato (*overvote*) o con votos considerados nulos (*undervote*)? ¿Describe con fidelidad esta relación el coeficiente de correlación calculado?
2. Parte del debate político actual acerca de la reforma del sistema de votación está en determinar si se deben requerir máquinas de voto con escáner óptico, para escanear las papeletas en las circunscripciones electorales (en presencia de los votantes) o en la oficina central (al concluir el día de las elecciones). ¿Cuál es el porcentaje medio de papeletas con votos a más de un candidato o considerados nulos en condados en los que se usa el escáner óptico en la oficina central, en comparación con los condados en los que se usa el escáner en todas las circunscripciones?

Para esta parte, entregue un archivo do. en STATA® y los datos en los que se reflejen las respuestas a estas preguntas. Escriba dos párrafos muy breves a partir de los resultados de STATA® para responder a estas preguntas.

Tenga en cuenta las siguientes pistas:

1. Necesitará hacer uso del comando "merge".
2. Una buena manera de hallar el número medio de "votos a más de un candidato" desglosado por condado y por el tipo de tecnología utilizada sería escribir en STATA:

table technology, c (mean overvote).

Si desea medir la participación de cada circunscripción en la media basándose en el número total de papeletas emitidas, escriba en STATA:

table technology [fweight=total\_ball], c(mean overvote).

Para saber más sobre el comando "table" y el uso de medidas, póngase en contacto con el profesor Hamilton o use el comando de ayuda "help" de STATA®.

Calcule la cantidad de tiempo que le ha llevado resolver este problema: \_\_\_\_ . \_\_\_\_ horas.

#### Datos de evaluación de la asignatura

Los siguientes datos (reales), recogidos al azar, provienen de la guía de evaluación de asignaturas de la primavera de 1995. Se pidió a los estudiantes que asignaran una puntuación, en una escala del 1 al 7 (1=pésimo, 7=excelente) la calidad global de la asignatura. El número de matrículas por asignatura se obtuvo a partir de los datos que figuran en la secretaría de la universidad.

ID	Puntuación	Número de matrículas
1	6	56
2	5,9	52
3	4,7	18
4	5,4	45
5	6,5	13
6	3,8	501
7	5,8	10
8	6	50
9	6	17
10	5,5	25
11	4,9	22
12	5,2	29
13	6	39
14	5,3	14
15	6,3	22
16	6,6	15

17	5,3	116
18	6,4	13
19	5,4	30
20	4,2	65
21	6,1	20
22	5,5	28
23	6,2	22
24	5	7
25	5,7	25
26	6,6	19

Variables en fla\_county\_subset.dta

Variable name	descripción	codificación	
county	nombre del condado	nombre natural del condado	
technology	tipo de tecnología de recuento de voto.	Datavote	papeletas que hay que perforar (sin marcar previamente)
		Manual	papeletas contadas a mano
		Palanca	máquinas de palanca mecánica
		Óptico	escáner óptico
		Votomatic	papeletas que hay (marcadas previamente)
centraltab	Papeletas pasadas por escáner óptico en la oficina central	1	circunscripción
		2	central
		3	no aplicable

Variables en fla\_precinct\_subset.dta

variable name	descripción
county	nombre del condado
precinct	nombre de la circunscripción/número
total_ball	número total de papeletas depositadas en la circunscripción

undervote	número total de paletas depositadas en la circunscripción que registran ausencia de voto para presidente (nota: no comprendo por qué pone que en el condado de Broward hay 1 papeleta <i>undervote</i> )
overvote	número total de papeletas depositadas en la circunscripción que registran más de un voto para presidente
blackrv	porcentaje de votantes afroamericanos registrados en la circunscripción