Homework 2

1 Suponga que Y tiene una distribución gamma con $\alpha = n/2$ para algún entero positivo n y β igual a algún valor especificado. Use el método de las funciones generadoras de momento para demostrar que $W = 2Y/\beta$ tiene una distribución χ^2 con n grados de libertad.

2 Suponga que la variable aleatoria Y tiene una distribución gamma con parámetros $\alpha = 2$ y β desconocida. En el Ejercicio 1 usted utilizó el método de las funciones generadoras de momento para demostrar un resultado general que implicaba que $2Y/\beta$ tiene una distribución χ^2 con 4 grados de libertad. Usando $2Y/\beta$ como cantidad pivote, deduzca un intervalo de confianza de 90% para β.

3 Suponga que la variable aleatoria Y es una observación de una distribución normal con media μ desconocida y varianza 1. Encuentre un

a intervalo de confianza de 95% para μ,

b límite de confianza superior de 95% para μ,

c límite de confianza inferior de 95% para μ.

4 Suponga que Y está distribuida normalmente con media 0 y varianza σ^2 desconocida. Entonces Y^2/σ^2 tiene una distribución χ^2 con grado de libertad 1. Use la cantidad pivote Y^2/σ^2 para hallar un

a intervalo de confianza de 95% para σ^2,

b límite de confianza superior de 95% para σ^2,

c límite de confianza inferior de 95% para σ^2.

5 Use las respuestas del Ejercicio 4 para hallar un

a intervalo de confianza de 95% para σ,

b límite de confianza superior de 95% para σ,

c límite de confianza inferior de 95% para σ.

6 Sea $Y_1, Y_2, \ldots, Y_n$ una muestra aleatoria de tamaño n de una población con distribución uniforme en el intervalo $(0, \theta)$. Sean $Y_{(n)} = \text{máx}(Y_1, Y_2, \ldots, Y_n)$ y $U = (1/\theta)Y_{(n)}$.

a Demuestre que U tiene función de distribución

$$F_U(u) = \begin{cases} 0, & u < 0, \\ u^n, & 0 \le u \le 1, \\ 1, & u > 1. \end{cases}$$

b Como la distribución de U no depende de θ, U es una cantidad pivote. Encuentre un límite de confianza inferior de 95% para θ.

7 Suponga que Y tiene la siguiente función de densidad de probabilidad

$$f_Y(y) = \begin{cases} \dfrac{2(\theta - y)}{\theta^2}, & 0 < y < \theta, \\ 0, & \text{en cualquier otro punto.} \end{cases}$$

a Demuestre que Y tiene función de distribución

$$F_Y(y) = \begin{cases} 0, & y \le 0, \\ \dfrac{2y}{\theta} - \dfrac{y^2}{\theta^2}, & 0 < y < \theta, \\ 1, & y \ge \theta. \end{cases}$$

b Demuestre que Y/θ es una cantidad pivote.

c Use la cantidad pivote del inciso b para hallar un límite de confianza inferior de 90% para θ.

8 Consulte el Ejercicio 7.

a Use la cantidad pivote del Ejercicio 7(b) para hallar un límite de confianza superior de 90% para θ.

b Si $\hat{\theta}_L$ es el límite de confianza inferior para θ obtenido en el Ejercicio 7(c) y $\hat{\theta}_U$ es el límite superior hallado en el inciso a, ¿cuál es el coeficiente de confianza del intervalo $(\hat{\theta}_L,\ \hat{\theta}_U)$?

9 Consulte el Ejemplo hecho en clase (exponencial) y suponga que Y es una sola observación de una distribución exponencial con media θ.

a Use el método de las funciones generadoras de momento para demostrar que $2Y/\theta$ es una cantidad pivote y tiene una distribución χ^2 con 2 grados de libertad.

b Use la cantidad pivote $2Y/\theta$ para deducir un intervalo de confianza de 90% para θ.

c Compare el intervalo que obtuvo en el inciso b con el intervalo obtenido en el Ejemplo hecho en clase.

10 Consulte el Ejercicio 9. Suponga que $Y_1, Y_2, \ldots, Y_n$ es una muestra de tamaño n para una distribución exponencial con media θ.

a Utilice el método de las funciones generadoras de momento para demostrar que $2\sum_{i=1}^{n} Y_i/\theta$ es una cantidad pivote y tiene una distribución χ^2 con $2n$ grados de libertad.

b Use la cantidad pivote $2\sum_{i=1}^{n} Y_i/\theta$ para deducir un intervalo de confianza de 95% para θ.

c Si una muestra de tamaño $n = 7$ da $\overline{y} = 4.77$, use el resultado del inciso b para dar un intervalo de confianza de 95% para θ.

11 Consulte los Ejercicios 2 y 10. Suponga que $Y_1, Y_2, \ldots, Y_n$ es una muestra de tamaño n de una población con distribución gamma con $\alpha = 2$ y β descono cida.

a Utilice el método de las funciones generadoras de momento para demostrar que $2\sum_{1}^{n} Y_i/\beta$ es una cantidad pivote y tiene una distribución χ^2 con $4n$ grados de libertad.

b Use la cantidad pivote $2\sum_{1}^{n} Y_i/\beta$ para deducir un intervalo de confianza de 95% para β.

c Si una muestra de tamaño $n = 5$ da $\overline{y} = 5.39$, use el resultado del inciso b para dar un intervalo de confianza para β.

12 Consulte el Ejercicio 11. Suponga que $Y_1, Y_2, \ldots, Y_n$ es una muestra de tamaño n de una población con distribución gamma con parámetros α y β.

a Si $\alpha = m$, donde m es un entero conocido y β es una incógnita, encuentre una cantidad pivote que tenga una distribución χ^2 con $m \times n$ grados de libertad. Use esta cantidad pivote para deducir un intervalo de confianza de $100(1 - \alpha)\%$ para β.

b Si $\alpha = c$, donde c es una constante conocida pero no un entero y β es una incógnita, encuentre una cantidad pivote que tenga una distribución gamma con parámetros $\alpha^\star = cn$ y $\beta^\star = 1$. Dé una fórmula para un intervalo de confianza $100(1 - \alpha)\%$ para β.

13 En este ejemplo, p_1 y p_2 se usaron para denotar las proporciones de refrigera-dores de las marcas A y B, respectivamente, que fallaron durante los períodos de garantía.

a En el nivel aproximado de 98% de confianza, ¿cuál es el mayor "valor creíble" para la diferencia en las proporciones de fallas de refrigeradores de las marcas A y B?

b En el nivel aproximado de 98% de confianza, ¿cuál es el menor "valor creíble" para la diferencia en las proporciones de fallas de refrigeradores de las marcas A y B?

c Si p_1 - p_2 es realmente igual a 0.2251, ¿cuál marca tiene la mayor proporción de fallas durante el período de garantía? ¿Qué tanto más grande?

d Si p_1 - p_2 es realmente igual a –0.1451, ¿cuál marca tiene la mayor proporción de fallas durante el período de garantía? ¿Qué tanto más grande?

e Como se observó en el Ejemplo 8.8, cero es un valor creíble de la diferencia. ¿Concluiría usted que hay evidencia de una diferencia en las proporciones de fallas (dentro del período de garantía) para las dos marcas de refrigeradores? ¿Por qué?

14 ¿Está menguando el romance de los estadounidenses con el cine? En una encuesta Gallup[1] de $n = 800$ adultos seleccionados aleatoriamente, 45% indicaron que el cine estaba mejorando mientras que 43% dijeron que el cine estaba empeorando.

a Encuentre un intervalo de confianza de 98% para p, la proporción total de adultos que dicen que el cine está mejorando.

b ¿El intervalo incluye el valor $p = 0.50$? ¿Piensa usted que una mayoría de adultos dice que el cine está mejorando?

15 Los administradores de un hospital deseaban estimar el número promedio de días necesarios para el tratamiento de enfermos internados entre las edades de 25 y 34 años. Una muestra aleatoria de 500 pacientes entre estas edades produjo una media y una desviación estándar igual a 5.4 y 3.1 días, respectivamente.

Construya un intervalo de confianza del 95% para la duración media de permanencia de la población de pacientes de la cual se extrajo la muestra.

16 Cuando se trata de anunciar, los "preadolescentes" no están listos para mensajes de línea dura que publicistas usan con frecuencia para llegar a los adolescentes. El estudio[2] del grupo Geppeto encontró que 78% de los "preadolescentes" entienden y disfrutan anuncios que son tontos por naturaleza. Suponga que el estudio comprendió $n = 1030$ "preadolescentes".

a Construya un intervalo de confianza de 90% para la proporción de "preadolescentes" que entienden y disfrutan anuncios que son tontos por naturaleza.

b ¿Piensa usted que "más de 75%" de todos los "preadolescentes" disfrutan anuncios que son tontos por naturaleza? ¿Por qué?

17 ¿Cuál *es* la temperatura corporal normal para personas sanas? Una muestra aleatoria de 130 temperaturas corporales en personas sanas proporcionadas por Allen Shoemaker[3] dio 98.25 grados y desviación estándar de 0.73 grados.

a Dé un intervalo de confianza de 99% para el promedio de temperatura corporal de personas sanas.

b El intervalo de confianza obtenido en el inciso a ¿contiene el valor de 98.6 grados, que es el promedio aceptado de temperatura citado por médicos y otros? ¿Qué puede usted concluir?

18 Una pequeña cantidad selenio, de 50 a 200 microgramos (μg) por día, es considerada esencial para una buena salud. Suponga que se seleccionaron muestras aleatorias independientes, de $n_1 = n_2 = 30$ adultos provenientes de dos regiones de Estados Unidos y que se registró para cada persona una ingesta diaria de selenio tanto de líquidos como de sólidos. La media y la desviación estándar de la ingesta diaria de selenio para los 30 adultos de la región 1 fueron $\overline{y} = 167.1$ μg y $s_1 = 24.3$ μg, respectivamente. Los estadísticos correspondientes para los 30 adultos de la región 2 fueron $\overline{y} = 140.9$ μg y $s_2 = 17.6$ μg. Encuentre un intervalo de confianza de 95% para la diferencia en la ingesta media de selenio para las dos regiones.

19 Los siguientes estadísticos son el resultado de un experimento realizado por P. I. Ward para investigar una teoría relativa al comportamiento de cambio de piel del macho *Gammarus pulex*, un pequeño crustáceo.[4] Si el macho cambia de piel mientras se aparea con una hembra, éste debe liberarla y perderla. La teoría es que el macho *Gammarux pulex* es capaz de posponer dicho cambio, con lo cual reduce la posibilidad de perder su pareja. Ward asignó aleatoriamente 100 parejas de machos y hembras a dos grupos de 50 cada uno. Las parejas del primer grupo se mantuvieron juntas (normal); las del segundo grupo fueron separadas. Se registró el tiempo de muda para machos y hembras, y las medias, desviaciones estándar y tamaños muestrales se ilustran en la tabla siguiente. (El número de crustáceos en cada una de las cuatro muestras es menor que 50 porque algunos en cada grupo no sobrevivieron hasta el tiempo de muda.). Ver la tabla 1.

a Encuentre un intervalo de confianza de 99% para la diferencia en tiempo promedio de muda para machos "normales" contra los "separados" de sus parejas.

b Interprete el intervalo.

1 *Fuente:* "Movie Mania Ebbing," Gallup Poll of 800 adults, http://www.usatoday.com/snapshot/news/2001-06-14-moviemania.htm., 16–18 de marzo de 2001

2. *Fuente:* "Caught in the Middle", *American Demographics*, julio de 2001, pp. 14-15.

3. *Fuente:* Allen L. Shoemaker, "What's Normal? Temperature, Gender and Heart Rate," *Journal of Statistics Education* (1966).

4. *Fuente:* "*Gammarus pulex* Control Their Moult Timing to Secure Mates," *Animal Behaviour* 32 (1984).

Tabla 1.

	Tiempo de muda (días)		
	Media	s	n
Machos			
Normal	24.8	7.1	34
Separados	21.3	8.1	41
Hembras			
Normal	8.6	4.8	45
Separados	11.6	5.6	48

20 A la mayoría de estadounidenses les gusta participar eventos deportivos o al menos verlos. Algunos sienten que los deportes tienen más que sólo valor de entretenimiento. En una encuesta de 1000 adultos, realizada por KRC Research & Consulting, 78% sintieron que los deportes de gran atractivo tienen un efecto positivo en la sociedad.[5]

a Encuentre un intervalo de confianza de 95% para el porcentaje del público que piensa que los deportes tienen un efecto positivo en la sociedad.

b La encuesta publicó un margen de error de "más o menos 3.1%". ¿Esto concuerda con la respuesta de usted al inciso a? ¿Qué valor de p produce el margen de error dado por la encuesta?

21 En una encuesta de la CNN/*USA Today*/Gallup a 1000 estadounidenses se les preguntó qué tan bien los describía el término *patriótico*.[6] Algunos resultados de la encuesta están contenidos en la siguiente tabla de resumen.

	Grupo de edad		
	Todos	18–34	60+
Muy bien	.53	.35	.77
Regular	.31	.41	.17
No muy bien	.10	.16	.04
Nada bien	.06	.08	.02

a Si los grupos de 18–34 y 60 o más años estaban formados de 340 y 150 individuos, respectivamente, encuentre un intervalo de confianza de 98% para la diferencia en proporciones de los individuos en estos grupos de edades que acordaron que *patriótico* los describía muy bien.

b Con base en el intervalo que obtuvo en el inciso a, ¿piensa que la diferencia en proporciones de los que se vieron a sí mismos como patrióticos es de hasta 0.6? Explique.

22 Para una comparación de los porcentajes de piezas defectuosas producidas por dos líneas de montaje, de cada línea se seleccionaron muestras aleatorias independientes de 100 piezas. La línea A produjo 18 piezas defectuosas en la muestra y la línea B contenía 12 piezas defectuosas.

a Encuentre un intervalo de confianza de 98% para la verdadera diferencia en proporciones de piezas defectuosas para las dos líneas.

b ¿Hay evidencia aquí que sugiera que una línea produce una proporción más alta de piezas defectuosas que la otra?

5. *Fuente:* Mike Tharp, "Ready, Set, Go. Why We Love Our Games—Sports Crazy," *U.S. News & World Report,* 15 de julio de 1997, p. 31.

6. *Fuente:* Adaptado de "I'am a Yankee Doodle Dandy," Kowledge Networks: 2000, *American Demographics,* julio de 2001, p. 9.

23 Sea Y una variable aleatoria binomial con parámetro p. Encuentre el tamaño muestral necesario para calcular p con tolerancia de no más de .05 con probabilidad .95 en las siguientes situaciones:

a Si se considera que p es aproximadamente .9.

b Si no se conoce información acerca de p (use $p = .5$ para calcular la varianza de $\hat{p}$).

24 Un servicio estatal de fauna silvestre desea calcular el número promedio de días que cada cazador con licencia se dedica a esta actividad realmente durante una estación determinada, con un límite en el error de estimación igual a 2 días de caza. Si los datos recolectados en estudios anteriores han demostrado que σ es aproximadamente igual a 10, ¿cuántos cazadores deben estar incluidos en el estudio?

25 Es frecuente que encuestadores por teléfono entrevisten entre 1000 y 1500 personas sobre sus opiniones en asuntos varios. ¿El rendimiento de los equipos de atletismo universitarios tiene un impacto positivo en la percepción del público del prestigio de las instituciones? Una nueva encuesta se va a efectuar para ver si hay diferencia entre las opiniones de hombres y mujeres sobre este asunto.

a Si se han de entrevistar 1000 hombres y 1000 mujeres, ¿con cuánta precisión podría usted estimar la diferencia en las proporciones que piensan que el rendimiento de sus equipos de atletismo tiene un impacto positivo en la percepción del público acerca del prestigio de las instituciones? Encuentre un límite para el error de estimación.

b Supongamos que usted estuviera diseñando la encuesta y desea estimar la diferencia en un par de proporciones, correcta a no más de .02, con probabilidad .9. ¿Cuántas entrevistas deben incluirse en cada muestra?

26 Consulte el Ejercicio 16. ¿Cuántos "preadolescentes" deben haber sido entrevistados para calcular la proporción de ellos que entienden y disfrutan de anuncios de naturaleza tonta, correcto a no más de .02, con probabilidad 0.99? Use la proporción del ejemplo anterior para calcular el error estándar de la estimación.

27 Supongamos que usted desea calcular el pH promedio de precipitaciones en una zona que sufre de fuerte contaminación debido a la descarga de humo de una planta de energía eléctrica. Suponga que σ está en las cercanías de .5 pH y que se desea que el cálculo varíe en no más de .1 de μ con probabilidad cercana a .95. ¿Aproximadamente cuántas precipitaciones deben estar incluidas en la muestra (una lectura de pH por precipitación)? ¿Sería válido seleccionar todos las muestras de agua de una sola precipitación? Explique.

28 Consulte el Ejercicio 27. Supongamos que usted desea calcular la diferencia entre la acidez media para lluvias en dos lugares diferentes, uno en una zona relativamente poco contaminada a lo largo del océano y la otra en una región sometida a fuerte contaminación del aire. Si se desea que el cálculo sea correcto al .1 pH más cercano con probabilidad cercana a .90, ¿aproximadamente cuántas precipitaciones (valores de pH) deben incluirse en cada muestra? (Suponga que la varianza de las mediciones de pH es alrededor de .25 en ambos lugares y que las muestras han de ser del mismo tamaño.)

29 Consulte el Ejercicio 22. ¿Cuántas piezas deben muestrearse de cada línea si un intervalo de confianza de 95% para la diferencia real entre las proporciones ha de tener un ancho de .2? Suponga que muestras de igual tamaño se tomarán de cada línea.

30 Consulte la comparación de la ingesta diaria de selenio para un adulto en dos diferentes regiones de Estados Unidos del Ejercicio 8.61. Suponga que se desea calcular la diferencia en la ingesta diaria promedio entre las dos regiones, con un error máximo de no más de 5 μg, con probabilidad .90. Si se planea seleccionar un número igual de adultos de las dos regiones (esto es, si $\mu_1 = \mu_2$), ¿qué tan grandes deben ser n_1 y n_2?