Simulación **Nombre:** Pauta

Código de honor:
No hemos dado ni recibido ayuda durante esta Tarea

Primavera 2019
Tarea 2

Fecha de Entrega: 02/10/18 **Firma** ______________

Este certamen contiene 11 páginas (incluyendo esta cubierta) y 8 preguntas. Cerciorece que su copia contiene todas las páginas. Ponga su iniciales arriba de cada página en el caso de que separe las hojas y estas se puedan perder.

Usted **PUEDE** utilizar una hoja A4 escrita en una de sus carillas para el certamen.

Se requiere que muestre su trabajo para cada problema en este certamen. Las siguientes reglas aplican:

- **Organize su trabajo**, de forma razonablemente ordenada, en el espacio entregado. Trabajo desorganizado difícil de evaluar recibirá poco o nada de puntaje (independiente de su exactitud).
- **Respuestas misteriosas o sin fundamentos no recibirán puntaje**. Una respuesta correcta, sin soporte de calculos, explicación, o trabajo algebraico **NO** recibirá puntaje; una respuesta incorrecta que sea el resultado de calculos intermedios correctos podría recibir puntaje parcial.
- Si necesita mas espacio, use el reverso de la página; indique claramente cuando haga esto.

No escriba en la tabla a la derecha.

Problem	Points	Score
1	10	
2	10	
3	15	
4	16	
5	16	
6	16	
7	15	
8	12	
Total:	110	

Probability theory

1. Una variable aleatoria Y tiene la siguiente función de densidad:

$$f_Y(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{3}{16}x^2 + \frac{1}{4} & \text{for } 0 \leq x < 2 \\ 0 & \text{for } 2 < x \end{cases}$$

(a) (4 points) ¿Cuál es el valor esperado de Y?

$$E[Y] = \int_{-\infty}^{\infty} a\, f(a)\, da$$

$$= \int_0^2 \frac{3}{16}a^3 + \frac{a}{4}\, da$$

$$= \left[\frac{3}{64}a^4 + \frac{a^2}{8}\right]_0^2$$

$$= 3/4 + 1/2 = 1.25$$

(b) (6 points) ¿Cuál es la función de densidad acumulada de Y? ¿Es más probable obtener un valor cercano a 1/2 o a 3/2?

$$F_Y(a) = \begin{cases} 0 , & a < 0 \\ \int_0^a 3/16\, b^2 + \frac{1}{4}\, db , & 0 \leq a < 2 \\ 1 , & a \geq 2 \end{cases}$$

$$= \begin{cases} 0 , & a < 0 \\ \left[\frac{1}{16}b^3 + \frac{1}{4}b\right]_0^a = \frac{a^3}{16} + \frac{a}{4} , & 0 \leq a < 2 \\ 1 , & a \geq 2 \end{cases}$$

2. Una máquina produce componentes que son ya sea aceptables o defectuosos. Después de observar 100 pares de componentes, la siguiente información fue recolectada:

		Segundo	
		Aceptable	Defectuoso
Primero	Aceptable	75	5
	Defectuoso	10	10

Dejémos que X_1 represente el primer componente y X_2 el segundo. Los trabajadores asignaron el valor 0 a un componente aceptable y el de 1 a un componente defectuoso. Conteste las siguientes preguntas:

(a) (4 points) Estime la función de probabilidad de masa conjunta para estas dos variables aleatorias.

$$f_{xy}(x,y) = \begin{cases} 0.75, & \text{if } x=A \text{ and } y=A \\ 0.1, & \text{if } x=D \text{ and } y=A \\ 0.05, & \text{if } x=A \text{ and } y=D \\ 0.1, & \text{if } x=D \text{ and } y=D \end{cases}$$

(b) (6 points) ¿Cuál es la probabilidad que el segundo componente sea defectuoso si el primer componente no es defectuoso? ¿Cuál es la probabilidad que el segundo componente sea defectuoso? ¿Son X_1 y X_2 independientes?

$$P(Y=D \mid X=A) = P(X=A \cap Y=D) / P(X=A)$$
$$= 0.05/0.8 = 0.0625$$

$$P(Y=D) = P(Y=D \cap X=D) + P(Y=D \cap X=A)$$
$$= 0.15$$

Cálculo de métricas en simulación

3. Usted decide testear si su entendimiento sobre como opera internamente una simulación es el adecuado. Para ello va a un restaurant y comienza a observar su operación y va registrando lo que ve en una hoja, tal como se observa en la Figura 1.

Evento	Hora	Descripción
Abre tienda	8:00:00	Se abre la tienda y el servidor queda desocupado
llega Cliente 1	8:12:26	Cliente 1 llega y comienza a ser atendido.
llega Cliente 2	8:21:36	Cliente 2 llega y espera en cola
Salida Cliente 1	8:25:30	Cliente 1 abandona la tienda y cliente 2 pasa a atención.
llega Cliente 3	8:27:00	Cliente 3 llega y espera en cola
Fin Observación	8:30:00	Fin periodo observación

Figure 1: Formulario de asignación de estudio de tiempos

(a) (3 points) ¿Cuál es el tiempo promedio en el sistema?

	Cliente 1	Cliente 2	Cliente 3
llega	8:12:26	8:21:36	8:27:00
sale	8:25:30	—	—
	13:04		

Tiempo promedio en el sistema
13:04

(b) (8 points) ¿Cuál es número promedio en el sistema?

Clientes	Periodo	# clientes * periodo (seg)
0	8:00:00 - 8:12:26	0
1	8:12:26 - 8:21:36	1 * 550
2	8:21:36 - 8:25:30	2 * 246
1	8:25:30 - 8:27:00	1 * 150
2	8:27:00 - 8:30:00	2 * 180
		1552

Promedio en sistema = $\frac{1552}{1800} = 0.86$ ← tiempo total

(c) (4 points) ¿Cuál es la utilización del servidor?

El servidor esta ocupado 1054 seg
(746 segundos desocupados)
La utilización es $1054/1800 = 0.59$

Validación de distribución

4. Usted recolectó 1000 observaciones de demanda de un determinado producto. Tal como aprendió en simulación hizo un histograma para observar como se agrupaba la información. Además usted sabe que debe ajustar una distribución teórica con tal de remplazar la información del histograma. Después de unos cuantos análisis usted ha llegado a la información de la Figura 2.

	BIN 1	BIN 2	BIN 3	BIN 4	BIN 5	BIN 6	BIN 7
# observado por BIN	130	200	250	140	190	65	25
densidad dist 1	0.09	0.27	0.25	0.15	0.2	0.06	0.04
densidad dist 2	0.13	0.16	0.21	0.19	0.16	0.09	0.06

Figure 2: Formulario de asignación de estudio de tiempos

(a) (9 points) Calcular el valor $\tilde{\chi}^2$ para cada una de las distribuciones.

1

$$\chi^2 = \frac{(130-1000*0.09)^2}{1000*0.09} + \frac{(200-1000*0.21)^2}{1000*0.21} + \frac{(250-1000*0.25)^2}{1000*0.25} +$$

$$\frac{(140-1000*0.15)^2}{1000*0.15} + \ldots + \frac{(25-1000*0.04)^2}{1000*0.04} = 25.462$$

2

$$\chi^2 = \frac{(130-1000*0.13)^2}{1000*0.13} + \frac{(200-1000*0.16)^2}{1000*0.16} + \ldots +$$

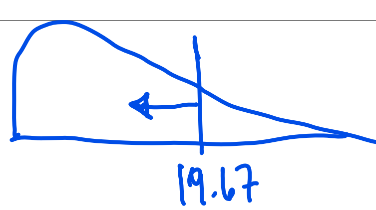
$$\frac{(25-1000*0.06)^2}{1000*0.06} = 63.76$$

(b) (4 points) Si el valor χ^2 de tabla es 19.67, ¿son adecuadas las distribuciones?

No son adecuadas ya que están en la zona de rechazo.

(c) (3 points) Un amigo le comenta que aumentando el nivel de significancia (α) usted podría tomar otra decisión. ¿Está su amigo en lo correcto?

No, dado que si aumento el nivel de significancia ocurre lo siguiente



lo que hace más difícil aceptar la hipótesis.

5. Usted quiere modelar el número diario de clientes que van a un restaurant. El dueño del local entrega una aproximación teórica de como los clientes se distribuyen durante la semana:

Día	L	M	M	J	V	S
Porcentaje (%)	15	10	15	20	25	15

Antes de utilizar esta distribución teórica usted decide validarla, por lo cual va a recolectar datos y obtiene la siguiente tabla:

Día	L	M	M	J	V	S	Total
Número de Clientes	26	18	34	45	50	27	200

(a) (16 points) Comprobar si la distribución teórica sirve para modelar el flujo de clientes

$$\chi^2_{calc} = \frac{(26-200*0.15)^2}{200*0.15} + \frac{(18-200*0.1)^2}{200*0.1} + \frac{(34-200*0.15)^2}{200*0.15} + \frac{(45-200*0.2)^2}{200*0.2} + \frac{(50-200*0.25)^2}{200*0.25} + \frac{(27*200*0.15)^2}{200*0.15}$$

$$\boxed{\chi^2_{calc} = 2.19}$$

Si $\chi^2_{calc} > \chi^2_{tabla}$ la distribución teórica no sirve, de lo contrario, es útil para modelar el número diario.

Likelihood y moment matching

6. Suponga que X es una variable aleatoria discreta con función de probabilidad de masa.

X	0	1	2	3
$P(X = x)$	$2\theta/3$	$\theta/3$	$2(1-\theta)/3$	$(1-\theta)/3$

(a) (16 points) Se han obtenido diez muestras de la distribución (3,0,2,1,3,2,1,0,2,1), y se le pide determinar el parámetro $\hat{\theta}$.

$$L(\theta|x) = \left(\frac{2\theta}{3}\right)^2\left(\frac{\theta}{3}\right)^3\left(\frac{2(1-\theta)}{3}\right)^3\left(\frac{(1-\theta)}{3}\right)^2$$

$$\ell(\theta|x) = 2\ln\left(\frac{2\theta}{3}\right) + 3\ln\left(\frac{\theta}{3}\right) + 3\ln\left(\frac{2(1-\theta)}{3}\right) + 2\ln\left(\frac{(1-\theta)}{3}\right)$$

$$\frac{d\,\ell(\theta|x)}{d\theta} = 2*\frac{3}{2\theta}*\frac{2}{3} + 3*\frac{3}{\theta}*\frac{1}{3} + 3*\frac{3}{2(1-\theta)}*\frac{-2}{3} + \frac{2*3}{(1-\theta)}*\frac{-1}{3}$$

$$= \frac{2}{\theta} + \frac{3}{\theta} - \frac{3}{(1-\theta)} - \frac{2}{(1-\theta)} = 0$$

$$\theta = 1-\theta$$

$$2\theta = 1$$

$$\boxed{\theta = \frac{1}{2}}$$

7. Se le pide determinar el estimador de "maximum likelihood" que caracteriza el parámetro óptimo de una distribución exponencial. La función de densidad se presenta a continuación.

$$f(x) = \begin{cases} \lambda_0 \exp(-\lambda_0 x_j) & \text{for } x \in [0, \infty) \\ 0 & \text{de otro modo} \end{cases}$$

(a) (12 points) Determine el parámetro $\hat{\lambda_0}$.

$$L(\lambda_0|x) = \prod_{j=1}^{n} \lambda_0 \exp(\lambda_0 x_j) = \lambda_0^n \exp(\lambda_0 \sum_{j=1}^{n} x_j)$$

$$\mathcal{L}(\lambda_0|x) = \ln(\lambda_0^n) + \ln(\exp(\lambda_0 \sum_{j=1}^{n} x_j))$$

$$\mathcal{L}(\lambda_0|x) = n \ln(\lambda_0) - \lambda_0 \sum_{j=1}^{n} x_j$$

$$\frac{d\,\mathcal{L}(\lambda_0|x)}{d\lambda_0} = \frac{n}{\lambda_0} - \sum_{j=1}^{n} x_j = 0$$

$$\lambda_0 = \frac{n}{\sum_{j=1}^{n} x_j} = \bar{X}^{-1}$$

(b) (3 points) ¿Que diferencia hay entre el parámetro calculado en (a) y el valor obtenido por medio del método de "Moment Matching"

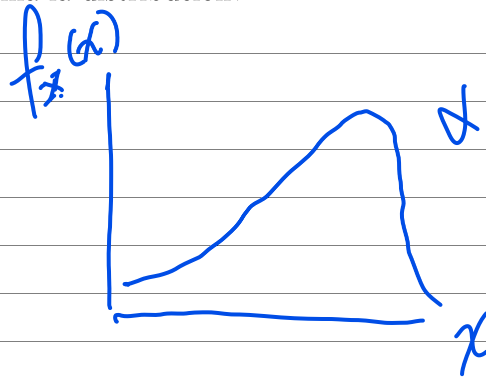
Técnicamente son el mismo parámetro y solo se observa una diferencia por como ha sido parametrizada la distribución.

Ajuste de distribuciones cuando no hay datos

8. En esta sección se evalúa su entendimiento del que hacer cuando no se dispone de datos.

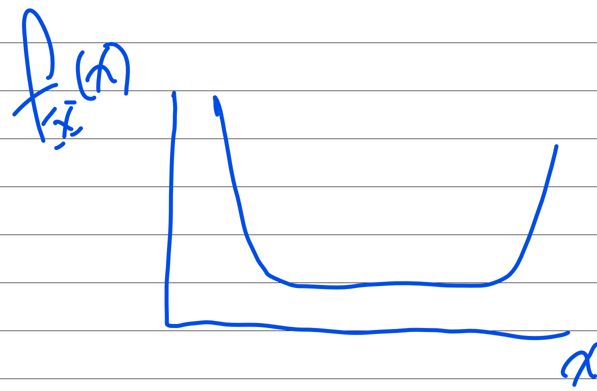
(a) (3 points) En una distribución Beta, si se cumple que $\alpha_1 > \alpha_2$ y ambos son mayores que

1, ¿qué forma toma la distribución?



$\alpha_1 > \alpha_2$

(b) (3 points) En una distribución Beta, si los parámetros α_1 y α_2 son ambos menores a 1, ¿qué forma toma la distribución?



$\alpha_1 < 1$ y

$\alpha_2 < 1$

(c) (6 points) Usted quiere modelar un proceso y no se mantienen registros de cuanto éste demora. Un operario le dice que como mínimo el proceso demora 5 minutos, que el valor mas observado por él es de 9 minutos, y que el valor más atípico que ha visto es 15 minutos. Ajuste una distribución Beta para este proceso.

$\hat{a} = 5$, $\hat{b} = 15$ $\hat{m} = 9$

$$r = \frac{\hat{b} - \hat{m}}{\hat{m} - \hat{a}} = \frac{15-9}{9-5} = 1.5$$

$$\alpha_1 = \frac{4 + 3r + r^2}{1 + r^2} = 3.31$$

$$\alpha_2 = \frac{1 + 3r + 4r^2}{1 + r^2} = 4.46$$

$$a + (b-a)\,\text{Beta}(\alpha_1, \alpha_2) = 5 + 10\,\text{Beta}(3.31, 4.46)$$