INFERENCIA ESTADÍSTICA
SOLEMNE
11-octubre-2018
Profesores: Jeanette Fuentes; Francisco Javier Leiva
Ayudantes: Sofía Muñoz; Francisca Quintanilla; Jaime Sánchez

I. Preguntas Conceptuales (20pts.)

1) En caso de tener dos estimadores para un parámetro, siempre será mejor escoger el estimador con menor varianza.
RESPUESTA:
Falso. Un estimador "a" podría tener menor varianza que un estimador "b", pero ser inconsistente, por lo cual no se acercaría en ningún caso al verdadero valor del parámetro que estamos buscando. El criterio de menor varianza sólo se aplica una vez que se ha verificado insesgadez o consistencia del estimador.

2) El teorema central del límite es una ley matemática que nos indica cuáles poblaciones se categorizan como finitas o infinitas.
RESPUESTA:
Falso, el teorema central del límite es un teorema que plantea que, en caso de muestras de observaciones de tamaño grande ($n > 30$)*, y que sean i.i.d, la media de la distribución tiende a una normal. Esto nada tiene que ver con la clasificación de las poblaciones entre finitas e infinitas para el cálculo de tamaño y error muestral.*

3) El tamaño de una muestra siempre debe ser proporcional al tamaño de la población, para así asegurar un nivel de error menor.
RESPUESTA:
Falso: El nivel de error depende de la varianza de la población, del nivel de confianza aceptado, y del tamaño de la muestra. Esto nada tiene que ver con el método de selección (que pudiera ser proporcional, aleatorio simple, por estratos, entre otros).

4) Dado que la función de distribución normal es simétrica, siempre su media tendrá el valor cero.
RESPUESTA:
Falso. Si bien la distribución normal es simétrica, esto es, su media es igual a su varianza, y su lado izquierdo es igual al derecho, ello no necesariamente lleva a que la media tenga un valor cero. El caso de una normal centrada en cero es un caso particular, pero esto no tiene por qué ser siempre así.

II. Estimadores (20pts.)

1) Dada la función de distribución: $f(x|k;\lambda) = x^k \cdot \exp\left(-\frac{x}{\lambda} - \lambda - \frac{k^2}{2}\right)$; estime, a través del método de máxima verosimilitud, los valores de k y de λ.
RESPUESTA:

$$f(x|k;\lambda) = x^k \cdot \exp\left(-\frac{x}{\lambda} - \lambda - \frac{k^2}{2}\right)$$
$$L(k;\lambda|x) = \prod_{i=1}^{n} x_i^k \cdot \exp\left(-\sum_{i=1}^{n}\frac{x_i}{\lambda} - \lambda n - \frac{k^2 n}{2}\right)$$
$$L(k;\lambda|x) = \left(\prod_{i=1}^{n} x_i\right)^k \cdot \exp\left(-\sum_{i=1}^{n}\frac{x_i}{\lambda} - \lambda n - \frac{k^2 n}{2}\right)$$
$$\ell(k;\lambda|x) = k \cdot \sum_{i=1}^{n} \ln(x_i) - \sum_{i=1}^{n}\frac{x_i}{\lambda} - \lambda n - \frac{k^2 n}{2}$$

Por lo tanto:

$$\frac{\partial \ell}{\partial k} = 0 \Leftrightarrow \sum_{i=1}^{n} \ln(x_i) - \frac{2kn}{2} = 0 \Rightarrow \hat{k}_{MV} = \frac{\sum_{i=1}^{n}\ln(x_i)}{n}$$
$$\frac{\partial \ell}{\partial \lambda} = 0 \Leftrightarrow +\frac{1}{\lambda^2}\sum_{i=1}^{n} x_i - n = 0 \Rightarrow \hat{\lambda}_{MV} = \left[\frac{\sum_{i=1}^{n} x_i}{n}\right]^{1/2}$$

2) Dada la función de distribución uniforme (continua): $f(x|a;b) = \frac{1}{b-a}$; estime, a través del método de momentos, los valores de *a* y *b*.
RESPUESTA:
Según lo que plantea el método de momentos los momentos muestrales serán igual a los momentos poblacionales, así del primer momento…

$$\frac{\sum_{i=1}^{n} x_i}{n} = E(x)$$
$$\Rightarrow \bar{x} = \frac{a+b}{2} \Rightarrow b = 2\bar{x} - a$$

Del segundo momento:

$$\frac{\sum_{i=1}^{n} x_i^2}{n} = E(x^2)$$
$$\Rightarrow \overline{x^2} = E(x^2)$$

Por definición:

$$E(x^2) - [E(x)]^2 = \text{var}(x)$$
$$\Rightarrow \overline{x^2} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}$$

$$\Rightarrow \overline{x^2} = \frac{a^2 + ab + b^2}{3}$$

Reemplazando:

$$\Rightarrow 3\overline{x^2} = a^2 + a(2\bar{x} - a) + 4(\bar{x})^2 - 4\bar{x}a + a^2$$
$$\Rightarrow 3\overline{x^2} = a^2 + 2\bar{x}a - a^2 + 4(\bar{x})^2 - 4\bar{x}a + a^2$$
$$\Rightarrow 0 = a^2 - 2\bar{x}a + 4(\bar{x})^2 - 3\overline{x^2}$$
$$\Rightarrow a = \frac{2\bar{x} \pm \sqrt{4(\bar{x})^2 - 16(\bar{x})^2 + 12\overline{x^2}}}{2}$$
$$\Rightarrow \hat{a} = \bar{x} - \sqrt{3\left[\overline{x^2} - (\bar{x})^2\right]}$$
$$\Rightarrow \hat{b} = \bar{x} + \sqrt{3\left[\overline{x^2} - (\bar{x})^2\right]}$$

III. Pregunta 3 (20pts.)

Un local comercial tiene el problema de que está atendiendo pocos clientes por día, por lo tanto, hemos sido contactados para solucionar el problema. Una vez tomadas una serie de medidas, hemos realizado una muestra de n = 34 días, en la cual el promedio de atención fue de 24,32 personas por día. Si se necesita que para que el negocio sea un éxito al menos deben ser atendidas 24 personas diarias… Responda lo siguiente: ¿Cuál es la probabilidad de que después de nuestra intervención nuestro negocio sea un éxito si la desviación estándar poblacional corresponde a σ = 5,05?

RESPUESTA:

Lo que queremos saber es la probabilidad de que el promedio sea mayor a 24, así:

$$\Pr(\bar{x} \geq 24) = 1 - \Pr(\bar{x} < 24)$$
$$\Pr(\bar{x} \geq 24) = 1 - \Pr\left(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} < \frac{24 - 24{,}32}{5{,}05/\sqrt{34}}\right)$$
$$\Pr(\bar{x} \geq 24) = 1 - \Pr(z < -0{,}37)$$
$$\Pr(\bar{x} \geq 24) = 1 - 0{,}3557$$
$$\Pr(\bar{x} \geq 24) = 0{,}6443$$

Por lo tanto, existe una probabilidad de 64,43% *de que el negocio tenga éxito.*

IV. STATA (20pts.)

A usted le presentan los resultados de una serie de cálculos realizados en STATA, hechos con una base de datos de un supermercado, la cual contiene las siguientes variables:

pasillo: Identifica con números del 1 al 4 los distintos pasillos del supermercado

codigo_item: Corresponde al código de cada uno de los productos vendidos en el supermercado

precio: Corresponde al precio en pesos ($) de cada producto.

pack: Variable que identifica si el producto se vende en forma individual o dentro de un "pack". Toma el valor 1 si "si" corresponde a un pack, y 2 si "no" corresponde a un pack.

promo: Variable que identifica si el producto está o no en promoción. Toma el valor 1 si "si" está en promoción y 2 si "no" está en promoción.

cantidad: corresponde al número de unidades de cada producto que hay disponible en el supermercado.

De acuerdo a lo anterior, responda las siguientes preguntas:

i) (2pts.) Indique si la variable codigo_item es una variable cuantitativa o cualitativa.
RESPUESTA:
La variable codigo_item es una variable cualitativa. Da lo mismo si el código es numérico, alfanumérico o gráfico, sólo representa una identificación (cualidad), no una medida.

ii) (3pts.) Indique qué representa la cifra de $4.960 presentada a continuación (indique claramente y en palabras qué es ese valor considerando exactamente las variables utilizadas).

```
. sum precio
```

Variable	Obs	Mean	Std. Dev.	Min	Max
precio	38	4960.263	4602.945	200	18990

RESPUESTA:
$4.960 *es el precio promedio de los artículos del supermercado.*

iii) (5pts.) Indique qué representan las cifras de $4.421 y de $5.152 presentadas más abajo.

```
. sum precio if pack==1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
precio	10	4421	3901.118	390	12690

```
. sum precio if pack==2
```

Variable	Obs	Mean	Std. Dev.	Min	Max
precio	28	5152.857	4880.071	200	18990

RESPUESTA:
$4.421 *es el precio promedio de todos los artículos que son vendidos dentro de un pack, y* $5.152 *es el precio promedio de todos los artículos que no son vendidos dentro de un pack.*

iv) (5pts.) Indique qué representan la variable **“var”** creada a continuación, y el porqué la suma del 100% de la variable contiene 16 observaciones y no 38 observaciones.

```
. gen var=1 if pack==1 & promo==1
(33 missing values generated)

. replace var=2 if pack==2 & promo==1
(11 real changes made)

. tab var
```

var	Freq.	Percent	Cum.
1	5	31.25	31.25
2	11	68.75	100.00
Total	16	100.00	

RESPUESTA:

Var es una variable dicotómica que identifica con el valor 1 a los artículos que están dentro de un pack y están en promoción, y con el valor 2 *a los artículos que no están dentro de un pack y están en promoción* (3*pts.*)

El total de observaciones de la variable es 16 *pues es una variable creada sólo para aquellos artículos que están en promoción. Aquellos que no están en promoción, quedarán con la celda vacía (missing value).* (2*pts.*)

v) (5pts.) Indique qué representa la variable llamada **“var3”** y por qué sólo contiene 4 observaciones.

```
. egen var2=max(precio), by(pasillo)

. gen var3=.
(38 missing values generated)

. replace var3=precio if precio==var2
(4 real changes made)

. tab var3
```

var3	Freq.	Percent	Cum.
8800	1	25.00	25.00
15490	1	25.00	50.00
17890	1	25.00	75.00
18990	1	25.00	100.00
Total	4	100.00	

RESPUESTA:

*La variable var*3 *representa el precio del artículo de mayor valor (precio) en cada uno de los pasillos.* (3*pts*).

La variable sólo contiene 4 observaciones porque el precio máximo se da en sólo 1 artículo por pasillo, y hay 4 *pasillos.* (2*pts*).

Formulario:

Bernoulli: $f(x|p) = p^x \cdot (1-p)^{1-x} \qquad \forall x = 0{,}1$

Binomial: $f(x|p) = \binom{T}{x} \cdot p^x \cdot (1-p)^{T-x} \qquad \forall x = 0{,}1{,}2, \ldots, T$

Poisson: $f(x|\lambda) = \frac{\lambda^x \cdot \exp(-\lambda)}{x!} \qquad \forall x = 0{,}1{,}2, \ldots$

Uniforme: $f(x|a,b) = \frac{1}{a-b} \qquad \forall a \leq x \leq b$

$\Rightarrow E(x) = \frac{a+b}{2}$

$\Rightarrow \mathrm{var}(x) = \frac{1}{12} \cdot (b-a)^2$

Exponencial: $f(x|\beta) = \frac{\exp(-x/\beta)}{\beta} \qquad \forall x \geq 0$

Normal: $f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot exp\left(-\frac{1}{2\sigma^2} \cdot (x-\mu)^2\right) \qquad \forall x \in \mathbb{R}$

$$\Pr(A_i|B) = \frac{\Pr(A_i) \cdot \Pr(B|A_i)}{\Pr(B)}$$

$$z_i = \frac{x_i - \mu}{\sigma} \qquad ; \qquad z_i = \frac{x_i - \mu}{\sigma/\sqrt{n}} \qquad ; \qquad z_i = \frac{x_i - \mu}{\sqrt{\frac{N-n}{N-1}}\left(\frac{\sigma}{\sqrt{n}}\right)}$$

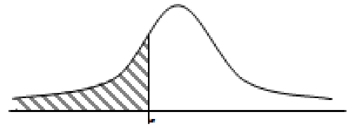
$$E(x) = \sum_{i=1}^{n} x_i \cdot f(x) \qquad ; \qquad E(x) = \int_{-\infty}^{+\infty} x_i \cdot f(x)\, dx$$

$$E(x^2) = \sum_{i=1}^{n} x_i^2 \cdot f(x) \qquad ; \qquad E(x^2) = \int_{-\infty}^{+\infty} x_i^2 \cdot f(x)\, dx$$

$$\mathrm{var}(x) = E(x^2) - [E(x)]^2$$

FUNCION DE DISTRIBUCION ACUMULATIVA NORMAL ESTANDAR

$$\Pr(Z \leq z) = F(z/0{,}1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp(-x^2/2)\,dx$$



	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
-3,5	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002	0,0002
-3,4	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002
-3,3	0,0005	0,0005	0,0005	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003
-3,2	0,0007	0,0007	0,0006	0,0006	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005
-3,1	0,0010	0,0009	0,0009	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007
-3,0	0,0013	0,0013	0,0013	0,0012	0,0012	0,0011	0,0011	0,0011	0,0010	0,0010
-2,9	0,0019	0,0018	0,0018	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
-2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
-2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
-2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
-2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
-2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
-2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
-2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
-2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
-2,0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
-1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
-1,8	0,0359	0,0351	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
-1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
-1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
-1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
-1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681
-1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
-1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
-1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
-1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
-0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
-0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
-0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
-0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
-0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
-0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
-0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
-0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
-0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998