Facultad de Ciencias — **Bioestadística**
Universidad de la Repúblicas — **Curso 2021**

Nombre:	**CI:**	**Año en que cursó**

Examen - 3 de Febrero de 2022

Problema 1: Para determinar si un determinado test es efectivo para detectar la enfermedad de Hansen, se realizó el test a una población de las cuales se sabe que el 3% tiene la enfermedad. El test detectó la enfermedad a un 98% que efectivamente la tenía y (erróneamente) a un 3% que no la tenía.

1 Calcular la probabildiad de que el test dé negativo a alguien que tiene la enfermedad.

2 Calcular la probabilidad de que el test dé positivo a una persona de dicha población.

3 Calcular la probabilidad de que la persona efectivamente esté enferma, dado que el test dió positivo.

Problema 2:
Queremos testear si efectivamente la proporción de personas de la población del ejercicio anterior que tienen la enfermedad de Hansen es del 3%. Para eso, realizamos el test a 200 personas y obtuvimos que 5 de ellas tenían la enfermedad. Usar que `qnorm(0.95)=1.645` y `qnorm(0.98)=2.054`.

1 Dar un valor estimado de la proporción de personas que tienen la enfermedad en base al dato que se obtuvo.

2 Calcular un intervalo de confianza al 96% para la proporción real de personas que tienen la enfermedad.

3 En base a los datos que se obtuvieron, realizar la prueba de hipótesis $H_0 : p = 0.03$ contra $H_1 : p = 0.05$, a nivel $\alpha = 0.05$.

Problema 3:
Consideremos la siguiente muestra de 10 datos,

$$-1.18 \quad 0.51 \quad -0.19 \quad 0.75 \quad 1.54 \quad 0.40 \quad -0.42 \quad 1.69 \quad 0.68 \quad -0.05.$$

Queremos ver si la misma puede considerarse, a un nivel de confianza de 95% como una muestra iid de una variable aleatoria.

1 Calcular el vector de rangos de la muestra anterior.

2 Plantear la prueba de hipótesis que se desea testear.

3 Usando la salida de `R`, indicar el valor del estadístico de Spearman de la muestra.

4 Usando la salida de `R`, indicar el p-valor de la prueba. Si se usa dicho p-valor, ¿qué concluye para la prueba planteada en 2?

```
> t= c(-1.18, 0.51, -0.19, 0.75, 1.54, 0.40, -0.42, 1.69,  0.68, -0.05)
> cor.test(t, sort(t), method = "spearman")

        Spearman's rank correlation rho

data:  t and sort(t)
S = 118, p-value = 0.4274
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.2848485
```

Solución

Problema 1: Sean los sucesos $E = \{\text{el paciente tiene la enfermedad}\}$, $+ = \{\text{el resultado del test es } +\}$ y $- = \{\text{el resultado del test es } -\}$. Se sabe que $\mathbb{P}(E) = 0.03$, $\mathbb{P}(+|E) = 0.98$ y $\mathbb{P}(+|E^c) = 0.03$.

1 $\mathbb{P}(-|E) = 1 - \mathbb{P}(+|E) = 1 - 0.98 = 0.02$.

2 $\mathbb{P}(+) = \mathbb{P}(+|E)\mathbb{P}(E) + \mathbb{P}(+|E^c)\mathbb{P}(E^c) = 0.98 \times 0.03 + 0.03 \times 0.97 = 0.0585$

3 $\mathbb{P}(E|+) = \frac{\mathbb{P}(+|E)\mathbb{P}(E)}{\mathbb{P}(+)} = \frac{0.98 \times 0.03}{0.0585} = 0.503$.

Problema 2: Sean $x_1, \dots, x_{200}$ tales que $x_i = \begin{cases} 1, & \text{si la persona } i \text{ está enferma,} \\ 0, & \text{si no} \end{cases}$.

1 Estimamos $p = \mathbb{P}(E)$ por $\hat{p} = \overline{x}_{200} = \frac{5}{200} = 0.025$.

2 $I_{1-\alpha} = [\overline{x}_{200} \pm k_\alpha]$ con $k_\alpha = \frac{\sigma}{\sqrt{200}} z_{1-\frac{\alpha}{2}}$. Estimamos $\sigma = \sqrt{p(1-p)}$ por $\hat{\sigma} = \sqrt{\overline{x}_{200}(1 - \overline{x}_{200})} = 0.156$ y usamos que $z_{1-\frac{\alpha}{2}} = z_{0.98} = 2.054$. Luego, $k_\alpha = 0.023$ y $I_{0.96} = [0.002, 0.048]$.

3 La región crítica para el test es $RC = \{\overline{X}_{200} > 0.03 + \frac{\hat{\sigma}}{\sqrt{n}} z_{1-\alpha}\} = \{\overline{X}_{200} > 0.048\}$. Como $\overline{x}_{200} = 0.025 < 0.048$, rechazamos H_0.

Problema 3:

1 $R = (1, 6, 3, 8, 9, 5, 2, 10, 7, 4)$

2 $\begin{cases} H_0: & \text{La muestra es i.i.d. ,} \\ H_1: & \text{NO} H_0 \end{cases}$.

3 $\rho_s = 1 - \frac{6S}{n(n^2-1)} = 0.2848485$ con $S = \sum_i (i - R_i)^2$

4 $p - valor = 0.4274$. Como $p - valor > \alpha = 0.05$, no se rechaza H_0.