Comentario: En todos los ejercicios propuestos

a) defina las variables aleatorias y los parámetros involucrados.

b) de ser posible indique:

   i. la distribución de las variables aleatorias

   ii. el significado intuitivo de los parámetros.

   (a) plantee las hipótesis nula y alternativa, e indique el nivel que usará para el test.

   (b) elija un test, calcule el valor del estadístico, calcule o acote el $p-$valor e indique la conclusión del test. Si el nivel del test no se especifica en el enunciado, tome por default 0.05.

   (c) compare los resultados de hacer las cuentas a mano con las salidas obtenidas con el R, de manera de chequear las primeras y aprender a usar las segundas, en aquellos ejercicios en los que ambas cosas sean posibles.

1. Una colonia de ratones de laboratorio tiene varios cientos de animales. El peso en gramos de los ratones adultos sigue una distribución normal con media igual a 30g. y desvío estándar de 5g. Como parte de un experimento, se les pidió a algunos estudiantes que eligieran 25 ratones adultos, sin ninguna premisa. El peso promedio de estos 25 animales fue de 33g. ¿Muestran estos datos evidencia suficiente a un nivel del 5% para decir que seleccionar los animales de esta manera no es lo mismo que elegirlos al azar? Justifique. Hallar el p-valor para el test elegido.

2. Consideremos un procedimiento para medir el contenido de manganeso en un mineral. A este procedimiento se lo ha usado muchas veces y se sabe que las mediciones siguen una distribución normal cuya desviación estándar es 0.15. Se está estudiando si el método tiene error sistemático.

   (a) Se hacen 8 determinaciones de un mineral preparado que tiene 7% de manganeso y se obtienen los siguientes resultados:

   6.90 7.10 7.20 7.07 7.15 7.04 7.18 6.95 $(\overline{x} = 7.074)$

   ¿Cuál es su conclusión si desea que la probabilidad de equivocarse al decir que el método tiene error sistemático cuando en realidad no lo tiene sea 0.05? (En este caso si no hay error sistemático las determinaciones siguen una distribución $N(7, 0.15^2)$ y si hay error la distribución es $N(\mu, 0.15^2)$ con $\mu \neq 7$).

   (b) Si el método tiene un error sistemático de 0.10 (o sea, si $\mu = 7.10$), ¿cuál es la probabilidad de cometer error de tipo II?

   (c) Se quiere aplicar un test estadístico de modo que, al igual que en el inciso a), la probabilidad de decir que hay error sistemático cuando no lo hay sea 0.05. Pero además se desea que si hay un error sistemático de 0.10, la probabilidad de detectarlo sea ≥ 0.80 (o lo que es equivalente, la probabilidad de error tipo II sea ≤ 0.20).

      i. El test del inciso a) ¿cumple con este requisito?

      ii. En caso contrario, ¿cuántas determinaciones habría que hacer como mínimo?

3. Una asociación de consumidores, preocupada por la cantidad de grasas contenida en una marca de hamburguesas, envía a un laboratorio independiente una muestra aleatoria de 12 hamburguesas para su análisis. El porcentaje de grasa en cada una de las hamburguesas de la muestra es:

$$21 \quad 18 \quad 19 \quad 16 \quad 18 \quad 24 \quad 22 \quad 19 \quad 24 \quad 14 \quad 18 \quad 15$$

El fabricante afirma que el contenido medio de grasa de este tipo de hamburguesas es menor al 18%. Basándose en la salida de R que figura más abajo, resuelva los siguientes items.

(a) La asociación de consumidores quiere saber si tiene motivos para decir que la afirmación del fabricante es falsa. Asumiendo que el contenido de grasa de cada hamburguesa de esta marca es una v.a con distribución normal y varianza deconocida, proponer un test para resolver este problema. Escribir las hipótesis a testear, definir las variables aleatorias y los parámetros involucrados en el test, escribir el estadístico del test y su distribución bajo la hipótesis nula y dar la región de rechazo. ¿Qué le informaría al representante de la asociación de consumidores como conclusión del test?

(b) Calcular el p-valor para el test del ítem (a). ¿Rechazaría la hipótesis nula a nivel 0.08?

(c) Intentar calcular la potencia del test en este caso cuando el contenido medio de grasa es 20%. Respecto del ejercicio anterior, ¿qué dificultad aparece cuando uno intenta hacer esta cuenta? Opcional: pensar alguna forma computacional de hacer esta cuenta.

```
hamburguesas<-scan()
21 18 19 16 18 24 22 19 24 14 18 15
c(mean(hamburguesas),var(hamburguesas),sd(hamburguesas))
[1] 19.000000 10.545455  3.247377
```

Sugerencia: usar que si $X_1, \ldots, X_n$ son i.i.d. $N(\mu, \sigma^2)$, entonces

$$\frac{\overline{X} - \mu}{s^2} \sim T_{n-1}$$

donde T_{n-1} es una T de Student con $n-1$ grados de libertad y $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$.

4. Se mide el grado de impurezas de un producto químico. El método de medición está afectado por un error que se supone $N(0, \sigma^2)$, σ^2 desconocido. Además los errores correspondientes a diferentes mediciones son independientes entre sí. Se hicieron 12 observaciones obteniendo que el promedio es 0.85 con un desvío estándar muestral de 0.05. A partir de estos datos, ¿hay evidencia significativa para decir que el grado de impurezas del producto es distinto de 0.7 a nivel 0.05? Hallar el p-valor

Sugerencia: usar que si $X_1, \ldots, X_n$ son i.i.d. $N(\mu, \sigma^2)$, entonces

$$\frac{\sum_{i=1}^{n}(X_i - \mu)^2}{\sigma^2} \sim \chi^2_n$$

donde χ^2_n es una Chi Cuadrado con n grados de libertad.