Maestría en Computo Estadístico
Inferencia Estadística
Tarea 9
13 de diciembre de 2020
Enrique Santibáñez Cortés
Repositorio de Git: Tarea 9, IE.

Potencia y tamaño de la muestra

Un artículo de 1992 en el Journal of the American Medical Association ('Una evaluación critíca de 98,6 grados F, el límite superior de la temperatura corporal normal y otros legados de Carl Reinhold August Wundrlich') informó la temperatura corporal, el género y la frecuencia cardíaca para una serie de temas. Las temperaturas corporales de 25 mujeres son las siguientes:

$$97{,}8, 97{,}2, 97{,}4, 97{,}6, 97{,}8, 97{,}9, 98{,}0, 98{,}0, 98{,}0, 98{,}1, 98{,}2, 98{,}3, 98{,}3, 98{,}4, 98{,}4, 98{,}4, 98{,}5, 98{,}6, 98{,}6, 98{,}7, 98{,}8, 98{,}8, 98{,}9, 98{,}9, y 99{,}0.$$

- La prueba de hipótesis es
$$H_0 : \mu = 98{,}6 \quad vs \quad H_A : \mu \neq 98{,}6$$
con $\alpha = 0{,}05$

1. Obtener la potencia si la verdadera temperatura corporal media femenina es tan baja como 98.

RESPUESTA
Tenemos que el error tipo II cuando se prueba con el juego de hipótesis bilateral es

$$\begin{aligned}\beta &= \mathbb{P}(-t_{\alpha/2,n-1} \leq T_{n-1} \leq t_{\alpha/2,n-1} | \delta \neq 0) \\ &= \mathbb{P}(-t_{\alpha/2,n-1} \leq T_{\delta,n-1} \leq t_{\alpha/2,n-1})\end{aligned}$$

donde $T_{\delta,n-1}$ denota la variable aleatoria no central t y $\delta = \frac{\mu_A-\mu_0}{\sigma/\sqrt{n}}$. Con lo anterior procedemos a calcular la potencia de la prueba cuando la alternativa $\mu_A = 98$ y considerando un nivel de significancia de $\alpha = 0{,}05$.

```r
# datos del problema
alpha <- 0.05
datos <- c(97.8, 97.2, 97.4, 97.6, 97.8, 97.9, 98.0, 98.0, 98.0, 98.1, 98.2, 98.3, 98.3,
           98.4, 98.4, 98.4, 98.5, 98.6, 98.6,
98.7, 98.8, 98.8, 98.9, 98.9, 99.0)

# tamaño de la m.a
n <- length(datos)
mu_0 <- 98.6
mu_A <- 98
d <- mu_0-mu_A

delta <- sqrt(n)*(mu_0-mu_A)/sd(datos)

# límites
t_inf <- -qt(p=1-alpha/2, df=n-1)
t_sup <- qt(p=1-alpha/2, df=n-1)

# calculamos el error
error_II <- pt(q=t_sup, df=n-1, ncp=delta)- pt(q=t_inf, df=n-1, ncp=delta)
```

Por lo tanto, tenemos que el error tipo II es

$$\beta = 3{,}1695396 \times 10^{-5},$$

y esto implica que la potencia es

$$\text{Potencia} = 0{,}9999683.$$

Por lo que podemos concluir que el tamaño de la muestra es el adecuado para proporcionar la sensibilidad deseada.

2. Modificar el código para poder graficar la potencia para $n = 2, 4, 15, 25, 40, 50$. dentro de un rango de $[0,\ 4]$, es decir $|\mu_A - \mu_0| < 4$. Comente lo que encontró.

RESPUESTA
Utilizando lo anterior, calculemos paras las n's y el rango de 4 de la siguiente forma

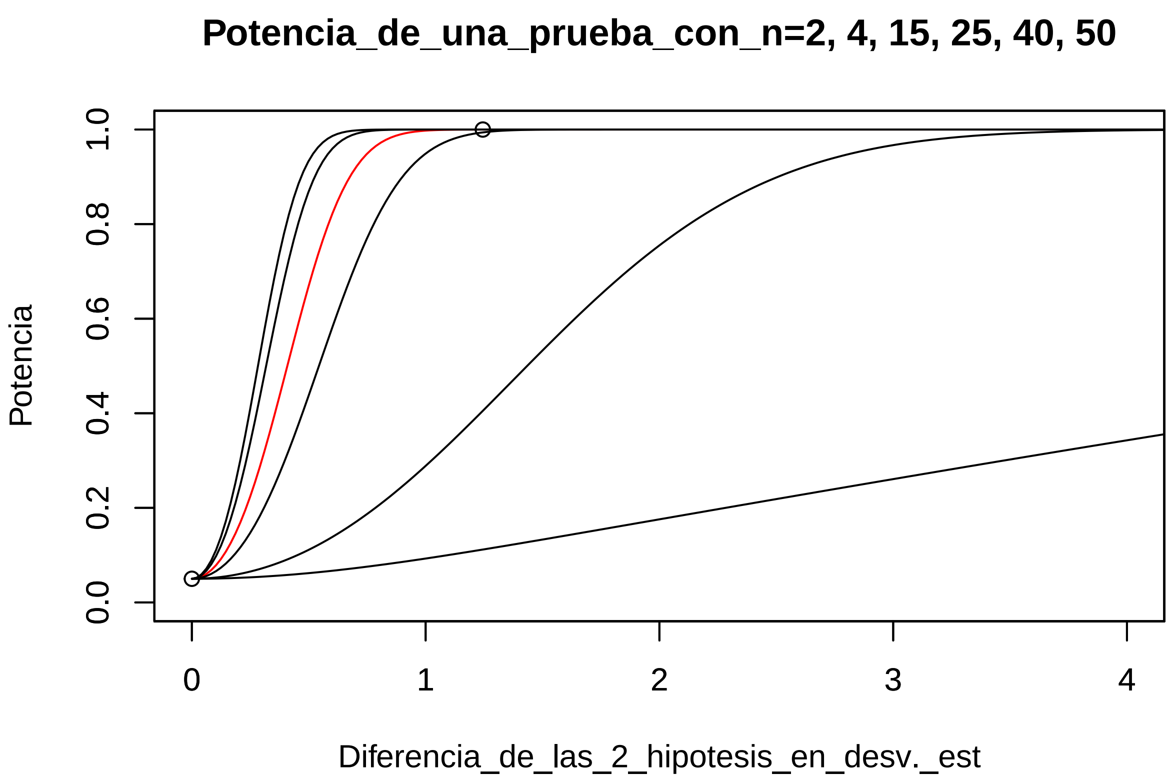
```r
# vector de distancias absolutas menores a 4.
rango_d <- seq(0, 4, 0.01)

# deltas
delta <- sqrt(n)*(rango_d)/sd(datos)

plot(rango_d/sd(datos),1-(pt(q=t_sup, df=n-1, ncp=delta)-pt(q=t_inf, df=n-1, ncp=delta)),
     type="l", ylab="", ylim=c(0,1),xlab="", xlim=c(0,4), col="red")
par(new=TRUE)
plot(d/sd(datos), 1-error_II, ylab="", ylim=c(0,1), xlab="", xlim=c(0,4),
     yaxt='n', xaxt='n')

size <- c(2, 4, 15, 40, 50)
for (i in size){ # Aquí se grafican las potencias para varios tamaños de muestra
  n2=i #Tamano de la muestra
  delta=sqrt(n2)*(rango_d)/sd(datos) #Vector de la delta para cada distancia
  par(new=TRUE)
    # límites
  t_inf <- -qt(p=1-alpha/2, df=n2-1)
  t_sup <- qt(p=1-alpha/2, df=n2-1)

  plot((rango_d)/sd(datos),1-(pt(q=t_sup, df=n2-1, ncp=delta)-pt(q=t_inf, df=n2-1, ncp=delta)),
       type="l", ylab="", ylim=c(0,1), xlab="", xlim=c(0,4), yaxt='n', xaxt='n')
}
par(new=TRUE)
plot(0,alpha,ylim=c(0,1),xlim=c(0,4),
ylab="Potencia",
xlab="Diferencia_de_las_2_hipotesis_en_desv._est",
main="Potencia_de_una_prueba_con_n=2, 4, 15, 25, 40, 50")
```



3. ¿Qué tamaño de muestra se necesitaría para detectar una temperatura corporal femenina media real tan baja como 98.4 si quisiéramos que la potencia de la prueba fuera de al menos 0.9?

HINT: Solo modificar beta en el código.

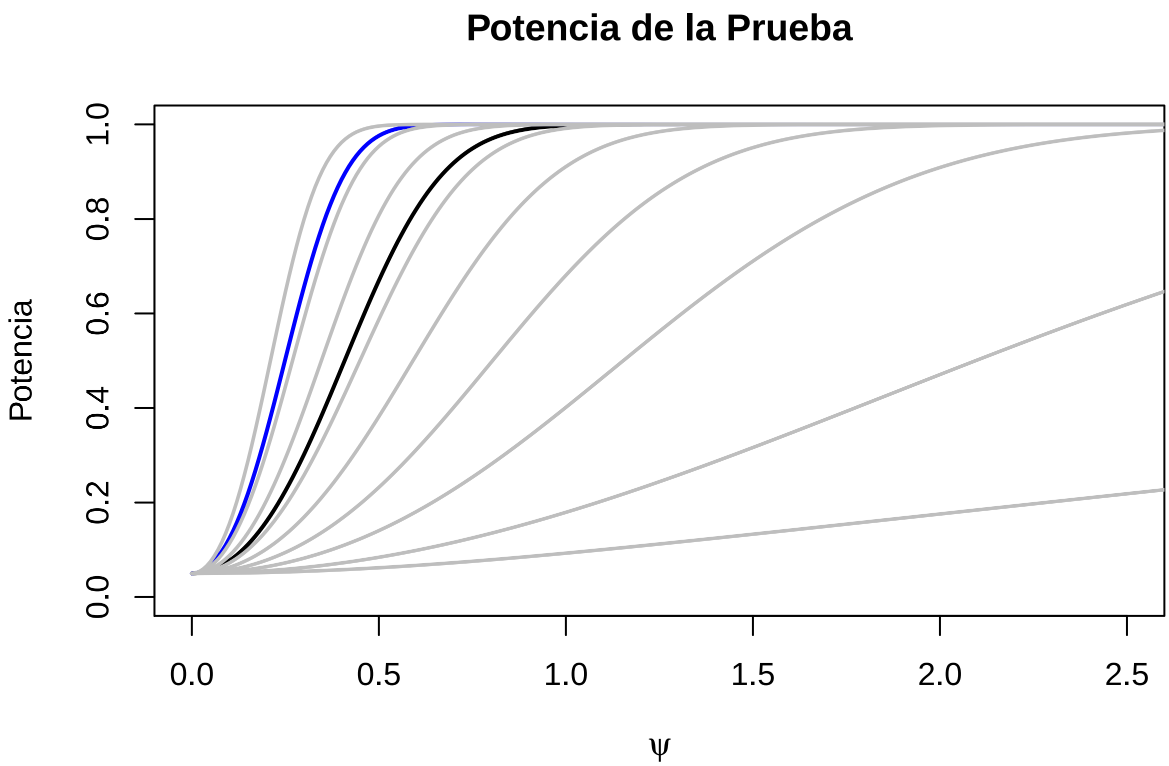
RESPUESTA
Modificando el código del ejemplo 2 para calcular el tamaño de muestra mínimo para obtener una potencia dada, tenemos que el código cambia de la siguiente forma.

```
fib <- function(n){ #Una función de Fibonacci para generar la muestra
  f = c(2,3)
  for(i in 3:n){
    f[i] = f[i-1] + f[i-2]
  }
  return(f)
}
minn <- function(psi,alf,potencia){ # Función para encontrar el tamaño
  for (i in 2:1000){
    # para una cierte distancia y una potencia
    x<-1-(pt(qt(1-alf,i-1),(i-1),sqrt(i)*psi)-pt(-qt(1-alf,i-1),(i-1),sqrt(i)*psi)) ## delta = sq
    if (x>potencia){
      return (i)
    }
  }
}
potenciaPSI <- function(alf,n,psi,p){
  cat("\n Prueba t para Una Muestra \n")
  cat("\n Tipo de Prueba: m = nula vs. mu diferente a nula \n")
  delta<-sqrt(n)*(psi)
  cat("\n Alpha =",alf)
```

```r
  minTM = minn(psi,alf,p)
  N <- fib(9)
  cat("\n n alternativas: \n")
  cat("\n",rev(N),"\n")
  cat("\n n minima para alcanzar la potencia (",p,"): \n")
  cat(" ",minTM,"\n")
  delta<-sqrt(n)*(psi) #El valor de delta
  t1<-qt(p=1-alf,df=n-1) #La t-student con alpha/2 y n-1 grados de libertad.
  beta<-pt(q=t1,df=n-1,ncp=delta)-pt(q=-t1,df=n-1,ncp=delta) # actualizamos beta
  potencia<-1-beta
  resultado = data.frame(Dif. = psi, taM = n, Pot = potencia)
  grafPot = plot(psi,1-beta,ylab="Potencia",xlab = expression(psi),
  ylim=c(0,1),xlim=c(0,2.5),main = "Potencia de la Prueba",
  col = "black",pch = 16,cex = 1.3)
  rangod = seq(0,4,0.01) # distancia de las media en desviaciones estandar
  delta = sqrt(n)*(rangod)
  grafPot = lines(rangod,1-(pt(q=t1,df=n-1,ncp=delta)-pt(q=-t1,df=n-1,ncp=delta)),lwd = 2)
  delta2 = sqrt(minTM)*(rangod)
  deltaTM = sqrt(minTM)*(psi)
  grafPot = lines(rangod,1-(pt(qt(1-alf,minTM-1),(minTM-1),delta2)-pt(-qt(1-alf,minTM-1),(minTM-1
  lwd = 2,col="blue")
  grafPot = points(psi,1-(pt(qt(1-alf,minTM-1),(minTM-1),deltaTM)-pt(-qt(1-alf,minTM-1),(minTM-1)
  pch = 16,cex = 1.3,col="blue")
  sapply(N, function(m) {
  delta = sqrt(m)*(rangod)
  grafPot = lines(rangod,1-(pt(qt(1-alf,m-1),(m-1),delta)-pt(-qt(1-alf,m-1),(m-1),delta)),
  col = "grey",lwd = 1.8)
  })
  RES = list(A = resultado, B = grafPot)
  return(RES)
}

potenciaPSI(0.05/2, 25, (98.4-98.6)/sd(datos), 0.9)
```

```
##
##  Prueba t para Una Muestra
##
##  Tipo de Prueba: m = nula vs. mu diferente a nula
##
##  Alpha = 0.025
##  n alternativas:
##
##  89 55 34 21 13 8 5 3 2
##
##  n minima para alcanzar la potencia ( 0.9 ):
##   64
```




```
## $A
##        Dif. taM       Pot
## 1 -0.4148699  25 0.5124379
##
## $B
## NULL
```

Por lo tanto, el tamaño de prueba mínimo que cumple con una potencia al menos de 0.9 es 64 ■.