1. **Actividad Teórica (Ejercicio 0.1):** Se pide demostrar el siguiente teorema:

*Teorema: Si el estadístico de prueba tiene una distribución continua, entonces bajo H0: θ = θ0, el p-valor tiene una distribución Uniforme [0,1].*

**Demostración:** Supongamos que D es el estadístico de Prueba para un contraste de hipótesis simple donde la hipótesis nula es *H0: θ = θ0* y que el estadístico de la prueba es continuo, cuya función de distribución bajo la hipótesis nula es FD(D)=Pr*H0*(D<d), donde d es el valor observado cuando una determinada muestra se considera.

Por definición sabemos que el P-valor = inf {α: D ∈ Rα}, donde la región de rechazo es Rα={D: D>K(α)} bajo la hipótesis simple considerada. Además para un valor D=d, se tiene un valor fijo del p-valor= PrH0(D>d)=1- FD(D).

De lo anterior, se puede concluir que el P-valor es una variable aleatoria ya que sus valores están en correspondencia uno a uno con los valores del estadístico D que es una variable aleatoria. En consecuencia podemos calcular Fp(p|H0) que es la función de distribución de P-valor bajo la hipótesis nula simple.

*Fp(p|H0)=PrH0(P<=p) (por definición de función de distribución)*

*=PrH0(1- FD(D)<=p) (por definición de P-valor cuando se tiene un p-valor fijo)*

*=PrH0(FD(D)>=1-p) (Despejando FD(D))*

*=PrH0(D>= F-1D(1-p)) (FD(D) es invertible, por continuidad del estadístico D)*

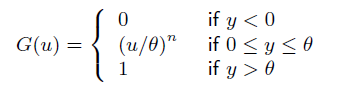
*= 1- FD (F-1D(1-p)) (Por definición de función de distribución)*

*=1-(1-p) (por propiedad de composición de funciones)*

*=p donde p ∈ (0, 1)*

Al derivar la función de distribución del P-valor obtenemos que su función de densidad es 1 y su soporte es p ∈ (0, 1). En consecuencia, el P-valor sigue una distribución *Uniforme [0,1] l.q.q.d.* En el script de R que se adjunta a este documento se ha desarrollado una simulación sobre el comportamiento del p-valor aplicando la simulación de Monte Carlo.

1. **Actividad Simulación (Ejercicio 0.2):** Sea *X1,X2,..Xn* una muestra i.id. de una distribución *Uniforme (0, θ)* y sea *D=max{ X1,X2,..Xn* *}.*
   1. **Probar que la función de distribución D es la siguiente:**



**Demostración:** En base a las condiciones del problema, se tiene que la distribución acumulada para *Xi* es *F(Xi)=* y por la independencia de las *Xi* la función de distribución conjunta para la muestra es igual a .

*Pr(D ≤ u) = Pr(max{ X1,X2,..Xn* *} ≤ u) = Pr(X1 ≤ u, . . . ,Xn ≤ u)*

*Pr(X1 ≤ u, . . . ,Xn ≤ u (Xi* *i.id.)*

*(Xi sigue una distribución Uniforme (0, θ))*

Por tanto, La distribución del estadístico *D=max{ X1,X2,..Xn* *},* es igual a *G(u).*

Considerando la siguiente prueba: H0: θ = 1 vs H1: θ > 1.

* 1. ¿Qué valor de *c*  hará que el nivel de significación de la prueba sea *0.05*?.

Por definición del nivel crítico, se tiene que *α = PrD(Rechazar Ho/ θ0)=PrD(D>c / θ0)=1-G(c)=1-.* Por tanto, despejando tenemos que *c= θ0 (1- α) (1/n)*

En consecuencia, con *α=0.05,* y bajo la hipótesis nulaθ=1*, tenemos que c=(0.95)(1/n)*

* 1. Para un nivel de significación *α* dado, encontrar la potencia en función de *θ* y *n*. Es la prueba propuesta consistente? Hacer algunos gráficos de la potencia como una función de *θ*, para varios valores de *n* y *α* (por ejemplo*, n = 10, 20, 50 y α = 0.05, 0.10).*

Dado que *c= θ0 (1- α) (1/n)* bajo la hipótesis nula a un nivel de significancia α, se tiene que la función de la potencia de la prueba paraθ > θ0 es*:*

*.* Por tanto, bajo la hipótesis nulaθ=1, se tiene que

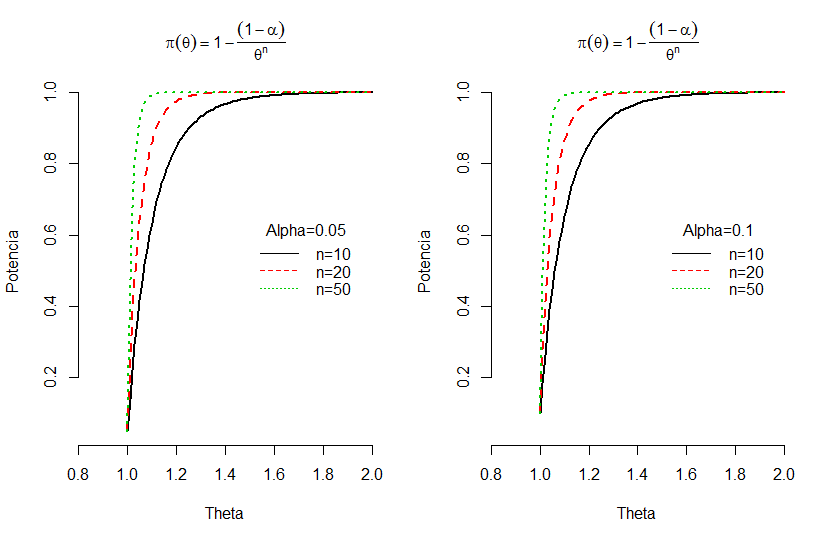
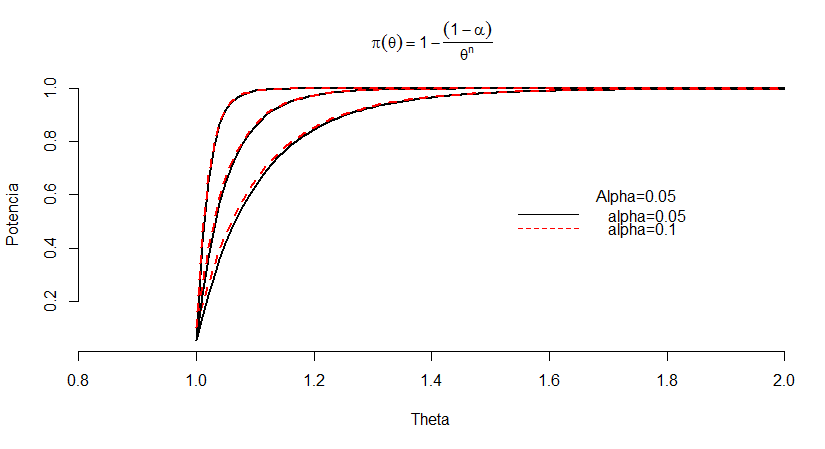


Figura 1: Gráficos de la potencia de prueba para n=10,20,50 y alpha=0.05,0.1

Claramente se puede observar en los gráficos de potencia, por un lado que a medida que aumenta el tamaño de la muestra tienden a 1 y al mismo tiempo cuando theta tiende a la hipótesis nula la potencia de la prueba tiende al valor de significancia fijado alpha. Por otro lado cuando se comparan las potencias de la prueba para diferentes valores de alpha se puede observar en la Figura 1 que prácticamente son las mismas funciones. En otras palabras no se cumple que la prueba sea UMP (Uniformily most powerful).

* 1. Prepare algunas simulaciones para estudiar el comportamiento del *p-valor*: En la Tabla 2, se presentan la estadística descriptiva del resultado de las simulaciones realizadas para los diferentes escenarios. Claramente en los resultados se puede observar que en el caso de que theta=1, los p-valores se distribuyen uniformemente en el intervalo [0, 1]. Por otro lado mientras theta va tomando valores mayores a la hipótesis nula, es decir para theta>1, la tendencia de los p-valores es a cero.

En la Figura 2, se presentan los respectivos gráficos de histograma y diagrama de cajas para los diferentes valores de theta cuando el tamaño de la muestra es 50. Los gráficos para los diferentes tamaños de muestra se los puede generar a partir del script en R que se adjunta a este trabajo.



Tabla 1: Estadística descriptiva de los p-valores simulados para los diferentes escenarios considerados

A partir de la Figura 2, para el caso donde theta=1, se tiene que se ha generado muestras tamaño 50 donde la hipótesis nula es cierta, dando como resultado de las simulaciones un histograma de los p-valores con una forma plana y uniforme sobre el intervalo [0, 1]. Por tanto podemos verificar que un p-valor no es la probabilidad de que Ho sea cierto. Esto es claro, ya que a pesar que Ho es cierta, el p-valor está uniformemente distribuido entre cero y uno.

Para los casos de theta>1, Se tiene que el histograma de los p-valores ya no es uniforme. Puede observarse que existe más chance de obtener p-valores menores al nivel de significación, además serán más altos estos valores bajo la hipótesis alternativa que bajo la hipótesis nula y ese efecto es más claro a medida que theta incrementa su valor.

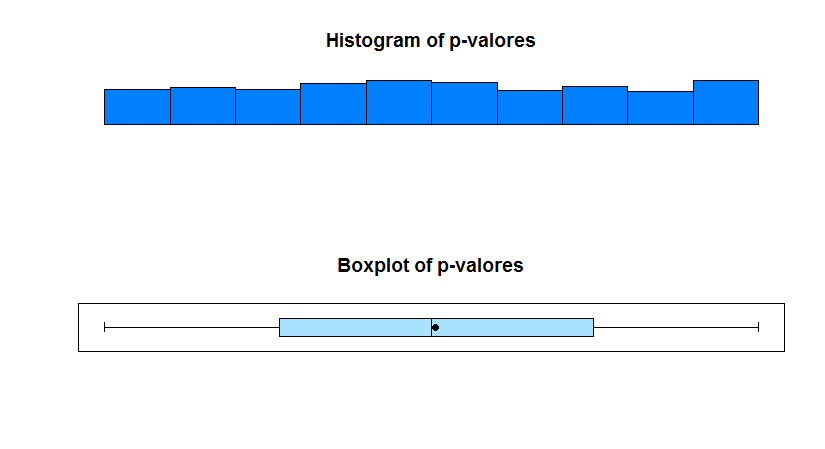
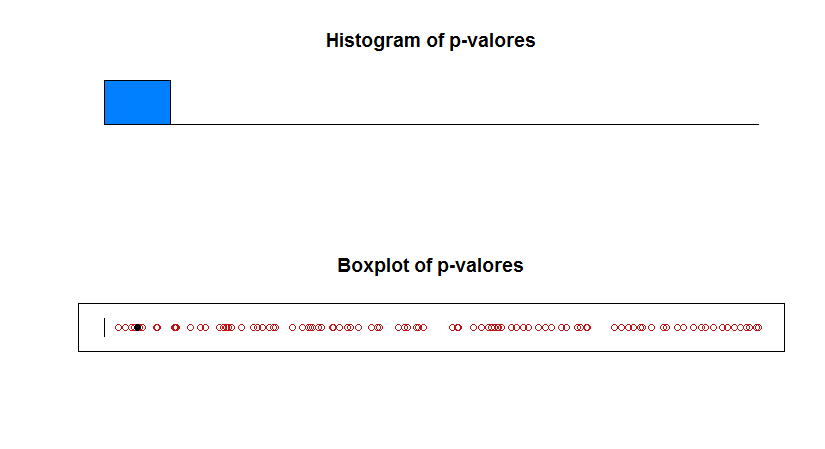
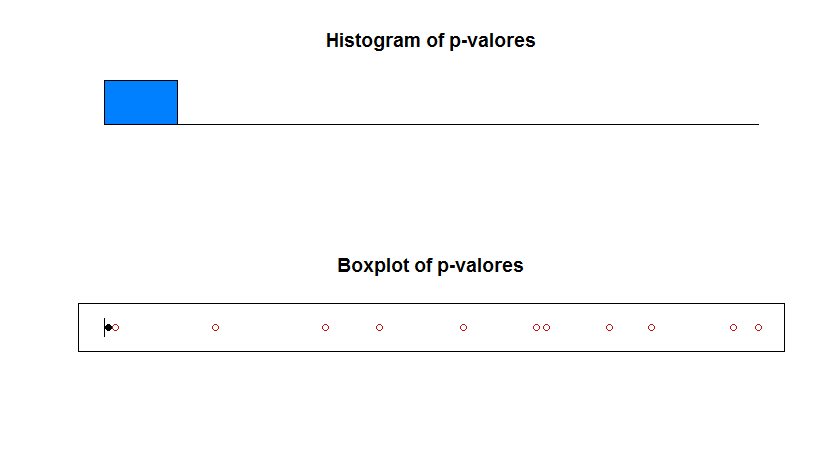


Figura 2: Histogramas y diagramas de cajas. De izquierda a derecha para los casos de theta=1, theta=1.05 y theta=1.1

1. **Bibliografía**

* Jun Shao (2003), Mathematical Statistics, second edition, Springer.
* Jun Shao (2005), Mathematical Statistics: Exercises and Solutions, Springer