TP Inferencia Estadística

Jesica Charaf e Ignacio Spiousas

25 de junio de 2023

El problema

Una fuente radiactiva emite partículas alfa de acuerdo con un proceso de Poisson de intensidad λ por segundo. Se tiene una fuerte sospecha de que el parámetro desconocido supera 0.5, y con el objetivo de confirmar dicha sospecha se medirán los tiempos entre emisiones consecutivas, obteniendo de esa manera una muestra de variables aleatorias independientes con distribución $\mathcal{E}(\lambda)$.

Se proponen dos alternativas, con respecto a la construcción del test de hipótesis:

1. Armar un test de nivel exacto α a partir del estadístico $T_1 = \sum_{i=1}^{n} X_i$.
2. Armar un test de nivel asintótico α a partir del estadístico $T_2 = \sqrt{n}\frac{\bar{X}-2}{2}$

El objetivo es comparar ambos métodos mediante una simulación en **R**, para luego elegir uno de ellos y realizar el experimento. Para cada valor de $n \in \{10, 30, 100, 1000\}$ se quiere estudiar el nivel empírico, y compararlo para las dos alternativas propuestas. También se desea aproximar y graficar la función potencia, y comparar.

Consideramos el siguiente procedimiento para la simulación:

- Generar una muestra de tamaño n de una distribución $\mathcal{E}(\frac{1}{2})$
- Calcular el valor del estadístico T para los dos test propuestos. Rechazar H_0 si el p-valor es menor al nivel 0.05.
- Repetir los ítems anteriores $N_{rep} = 10000$ veces, contando el total de veces en que se rechazó H_0. Se puede estimar el nivel de este test como la proporción empírica de rechazos, es decir, dividiendo la cantidad total de rechazos por N_{rep}, para cada valor de n, y armar una tabla con los resultados obtenidos.

Se pide:

a. Plantear claramente las hipótesis.

b. Justificar la elección de los dos estadísticos que propone el enunciado, y determinar sus distribuciones (exacta o asintótica) bajo H_0.

c. Realizar la simulación pedida para los diferentes valores de n, calcular el nivel empírico y expresarlo en una tabla.

d. Para cada uno de los valores propuestos de n y eligiendo una grilla de valores para λ, aproximar y graficar la función potencia.

e. Concluir en base a los resultados observados.

La resolución

a. Plantear claramente las hipótesis.

Consideramos que lo razonable es que la hipótesis nula sea que el parámetro **NO** supera el valor de 0.5 y la hipótesis alternativa que **SÍ** lo supera. Como dice el enunciado, se tiene la "sospecha" de que es mayor, por ende, el *status quo* es que el parámetro es menor. Simbólicamente:

$$H_0 : \lambda \leq 0.5 \qquad y \qquad H_1 : \lambda > 0.5$$

b. Justificar la elección de los dos estadísticos que propone el enunciado, y determinar sus distribuciones (exacta o asintótica) bajo H_0.

Primero vamos a desarrollar el **test de nivel exacto** α.

Partimos de que las mediciones son i.i.d y tienen la siguiente distribución:

$$X_1, ..., X_n \overset{i.i.d}{\sim} \mathcal{E}(\lambda)$$

O sea que la función de densidad de cada X es:

$$f_X(x) = \lambda e^{-\lambda x} \mathcal{I}_{\{x \geq 0\}}$$

Justamente por ser independientes, la densidad condicional se puede escribir como la productoria de las densidades individuales

$$f(\underset{\sim}{x}) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} \mathcal{I}_{\{x_i \geq 0\}}$$

Y por ser idénticamente distribuidas podemos operar y llegar a:

$$f(\underset{\sim}{x}) = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i} \prod_{i=1}^{n} \mathcal{I}_{\{x_i \geq 0\}}$$

De este resultado podemos ver que se trata de una distribución de familia exponencial donde:

$$A(\lambda) = \lambda^n$$
$$C(\lambda) = -\lambda$$
$$r(\underset{\sim}{x}) = \sum_{i=1}^{n} x_i$$
$$h(\underset{\sim}{x}) = \prod_{i=1}^{n} \mathcal{I}_{\{x_i \geq 0\}}$$

Entonces el estadístico (T_1 desde ahora) es igual $r(\underset{\sim}{X})$, de manera que el mismo queda definido como:

$$T_1 = r(\underset{\sim}{X})$$

$$T_1 = \sum_{i=1}^{n} X_i$$

Recuperando el estadístico propuesto en el enunciado del problema y teniendo en cuenta que $C(\lambda)$ es una función decreciente con λ, estamos en condiciones de escribir el test uniformemente más potente (UMP) de la siguiente forma:

$$\delta(\underset{\sim}{X}) = \begin{cases} 1 & si \quad -\sum_{i=1}^{n} X_i > K_\alpha \\ 0 & si \quad no \end{cases} \quad ,$$

con K_α tal que:

$$P_{\lambda=0.5}[\delta(\underset{\sim}{X}) = 1] = \alpha.$$

Entonces, lo siguiente que tenemos que hacer es calcular K_α, y para esto vemos a "pararnos" en un mundo en el que H_0 es verdadera. Bajo el mundo H_0, vamos a considerar al parámetro como $\lambda = 0.5$ y, por lo tanto, la distribución de cada una de las variables aleatorias será $X_i \overset{i.i.d}{\sim} \mathcal{E}(\frac{1}{2})$. Con esto en mente, podemos determinar la distribución de la suma de variables aleatorias (que es el estadístico T_1):

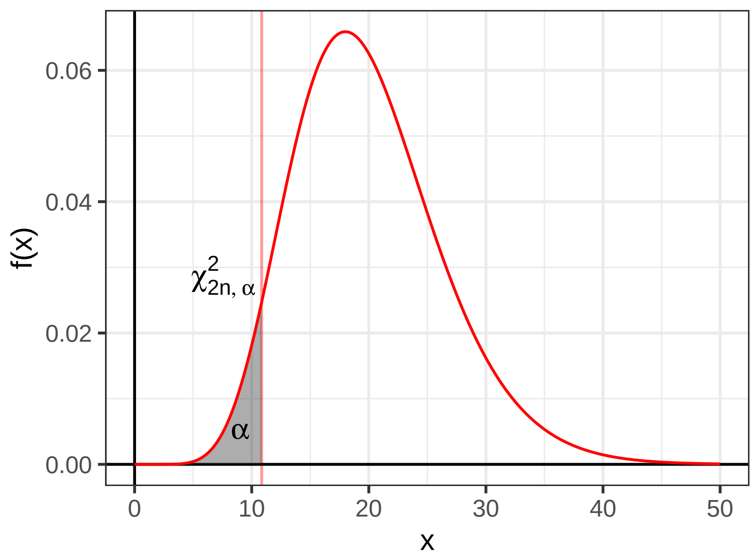
$$\begin{aligned} \sum_{i=1}^{n} X_i &\sim \Gamma\left(n, \frac{1}{2}\right) \\ &\sim \chi^2_{2n} \end{aligned}$$

Es decir, la suma de variables aleatorias $X_i \overset{i.i.d}{\sim} \mathcal{E}(\frac{1}{2})$ tiene distribución χ^2_{2n}. Entonces ahora vayamos a despejar K_α de la condición $P_{\lambda=0.5}[\delta(\underset{\sim}{X}) = 1] = \alpha$

$$\begin{aligned} \alpha &= P_{\lambda=0.5}[\delta(\underset{\sim}{X}) = 1] \\ &= P_{\lambda=0.5}(-\sum_{i=1}^{n} X_i > K_\alpha) \\ &= P_{\lambda=0.5}(\sum_{i=1}^{n} X_i < -K_\alpha) \\ &= P_{\lambda=0.5}(\sum_{i=1}^{n} X_i < \chi^2_{2n,\alpha}) \end{aligned}$$

Es decir, $-K_\alpha$ es igual al cuantil α de la distribución Chi cuadrado de grados de libertad $2n$ ($\chi^2_{2n,\alpha}$).

Por ejemplo, en la figura que se muestra a continuación podemos ver, para $n = 10$ y $\alpha = 0.05$, el cuantil de la distribución (línea roja vertical, $\chi^2_{2n,\alpha}$) junto con el área gris que representa a la probabilidad acumulada de $\alpha = 0.05$.



Finalmente, podemos definir al test de nivel exacto como:

$$\delta(\underset{\sim}{x}) = \begin{cases} 1 & si \quad \sum_{i=1}^{n} X_i < \chi^2_{2n,\alpha} \\ 0 & si \quad no \end{cases} \quad .$$

Ahora sigamos con el **test de nivel asintótico** α.

Partimos de nuevo de que las mediciones son i.i.d y tienen la siguiente distribución:

$$X_1, ..., X_n \overset{i.i.d}{\sim} \mathcal{E}(\lambda)$$

Sabemos que $E(X_i) = \frac{1}{\lambda}$ y que $V(X_i) = \frac{1}{\lambda^2}$, por lo tanto $E(\bar{X}) = \frac{1}{\lambda}$ y que $V(\bar{X}) = \frac{1}{n}\frac{1}{\lambda^2}$. Entonces, usando el teorema central del límite podemos decir que:

$$\frac{\bar{X} - \frac{1}{\lambda}}{\sqrt{\frac{1}{n\lambda^2}}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$$

$$\sqrt{n}\frac{\bar{X} - \frac{1}{\lambda}}{1/\lambda} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$$

Ahora volvemos a "pararnos" en H_0 y a considerar a $\lambda = 0.5$ (el supremo de Θ_1) y nos queda:

$$\sqrt{n}\frac{\bar{X} - \frac{1}{0.5}}{1/0.5} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$$

$$\sqrt{n}\frac{\bar{X} - 2}{2} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)\,,$$

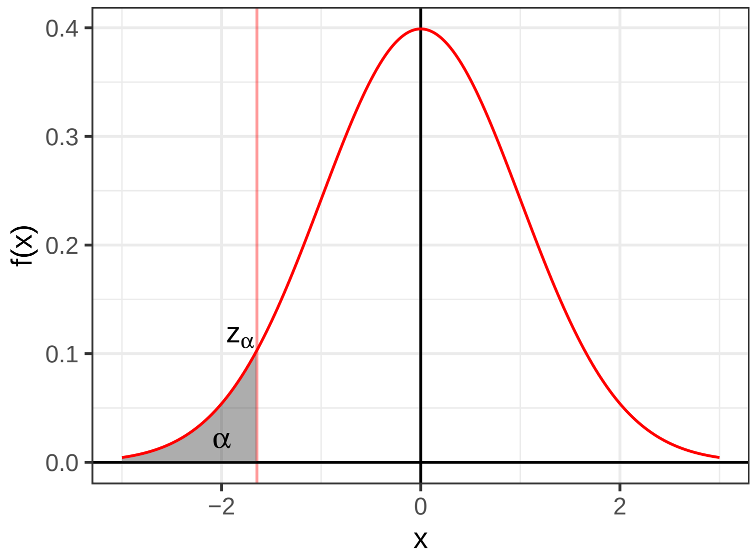
que corresponde con el estadístico T_2 planteado en el enunciado. Luego, hallamos un K'_α tal que:

$$P_{\lambda=0.5}\left(\sqrt{n}\frac{\bar{X} - 2}{2} < K'_\alpha\right) \overset{n\to\infty}{\to} \alpha \quad .$$

De forma que el test asintótico de nivel α queda definido como:

$$\delta(\underset{\sim}{x}) = \begin{cases} 1 & si \quad \sqrt{n}\frac{\bar{X}-2}{2} < Z_\alpha \\ 0 & si \quad no \end{cases} \quad .$$

En este caso el cuantil que tenemos que buscar es el de una distribución $\mathcal{N}(0,1)$. Por ejemplo, para $\alpha = 0.05$:



c. Realizar la simulación pedida para los diferentes valores de n, calcular el nivel empírico y expresarlo en una tabla.

Ahora lo que vamos a hacer es simular el "experimento", generando muestras de tamaño n con distribución $\mathcal{E}(\frac{1}{2})$, y medir la cantidad de veces que se rechazó H_0 para ver cómo se relaciona eso con el n para cada uno de los tests propuestos (el UMP y el asintótico). Vamos a realizar `Nrep=10000` simulaciones de cada test con valores de `n = c(10,30,100,1000)`.

```
# La función simulación simula el experimento y se fija el resultado del test
simulacion_UMP <- function(Nrep, n, lambda){
  rechazos <- rep(NA,Nrep)
  for (i in 1:Nrep) {
    set.seed(2*i) # Para tener la misma muestra para cada simulación
    muestra <- rexp(n, lambda)
    estadistico <- sum(muestra) # El estadístico del test UMP
    rechazos[i] <- pchisq(estadistico, 2*n)<0.05 # El test compara con Chisq
  }

  alpha_emp <- mean(rechazos)
  return(alpha_emp)
}

simulacion_asin <- function(Nrep, n, lambda){
  rechazos <- rep(NA,Nrep)
  for (i in 1:Nrep) {
    set.seed(2*i) # Para tener la misma muestra para cada simulación
    muestra <- rexp(n, lambda)
    estadistico <- sqrt(n)*(mean(muestra)-2)/2 # El estadístico del test asintótico
    rechazos[i] <- pnorm(estadistico, 0, 1)<0.05 # El test compara con Norm
```

```r
  }

  alpha_emp <- mean(rechazos)
  return(alpha_emp)
}

set.seed(123)
ns <- c(10,30,100,1000)

tabla <- tibble(n = ns) %>%
  rowwise() %>%
  mutate(alpha_emp_UMP = simulacion_UMP(Nrep = 10000, n = n, lambda = 0.5),
         alpha_emp_asin = simulacion_asin(Nrep = 10000, n = n, lambda = 0.5))

tabla %>%
  knitr::kable(col.names = c("n", "alfa empírico UMP", "alfa empírico asintótico"))
```

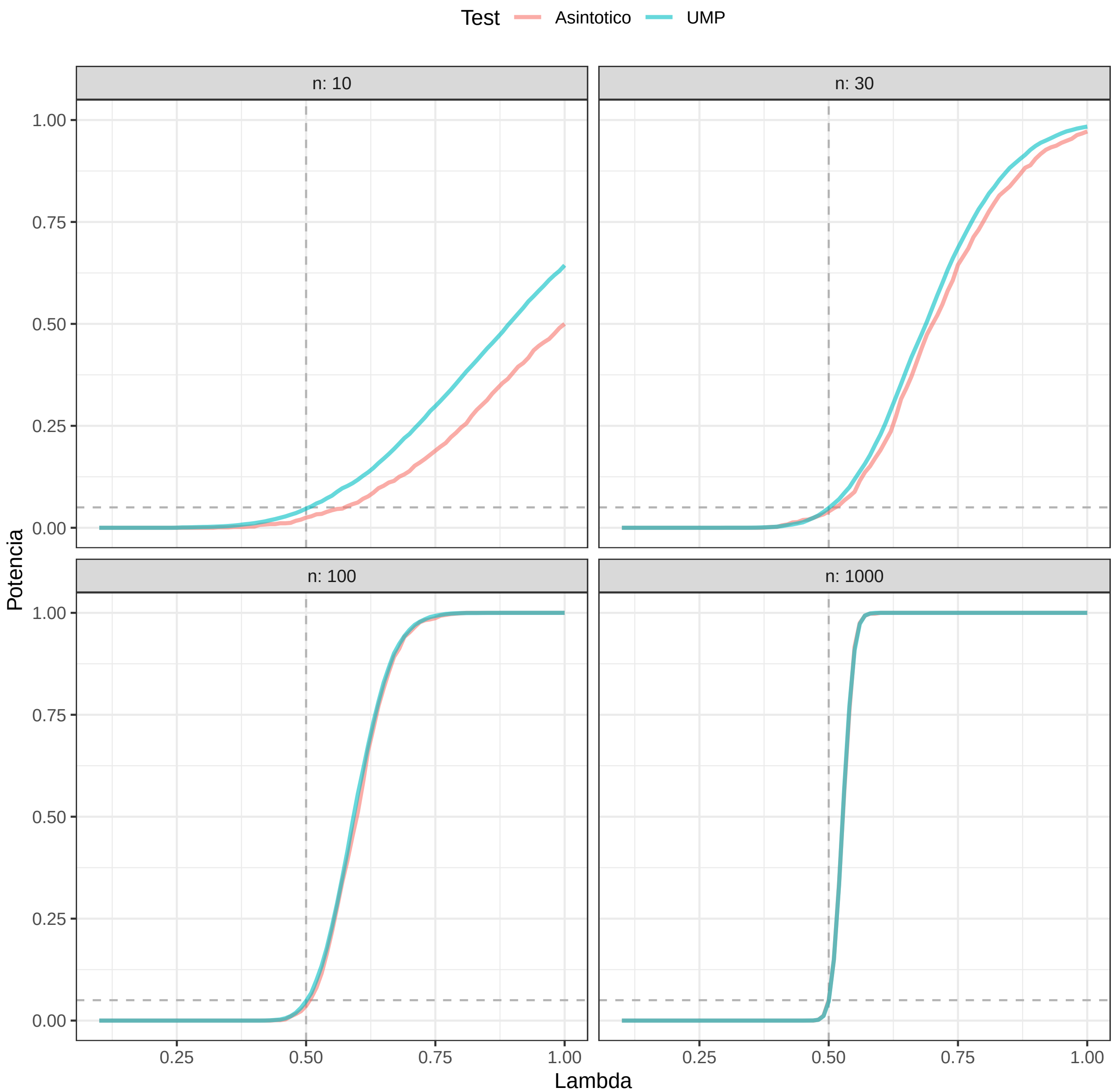
n	alfa empírico UMP	alfa empírico asintótico
10	0.0470	0.0224
30	0.0483	0.0353
100	0.0477	0.0394
1000	0.0469	0.0451

d. Para cada uno de los valores propuestos de n y eligiendo una grilla de valores para λ, aproximar y graficar la función potencia.

Para estas simulaciones vamos a utilizar el mismo `Nrep` y los mismos valores de `n` del inciso anterior pero nos vamos a mover del $\lambda = 0.5$ y vamos a barrer valores de lambda desde 0.1 a 1, con un paso de 0.01. De esta manera, para los distintos valores de λ generaremos muestras con distribución $\mathcal{E}(\lambda)$ y estimaremos la probabiliad de rechazar H_0 (es decir, la potencia) a partir de la proporción empírica de rechazos.

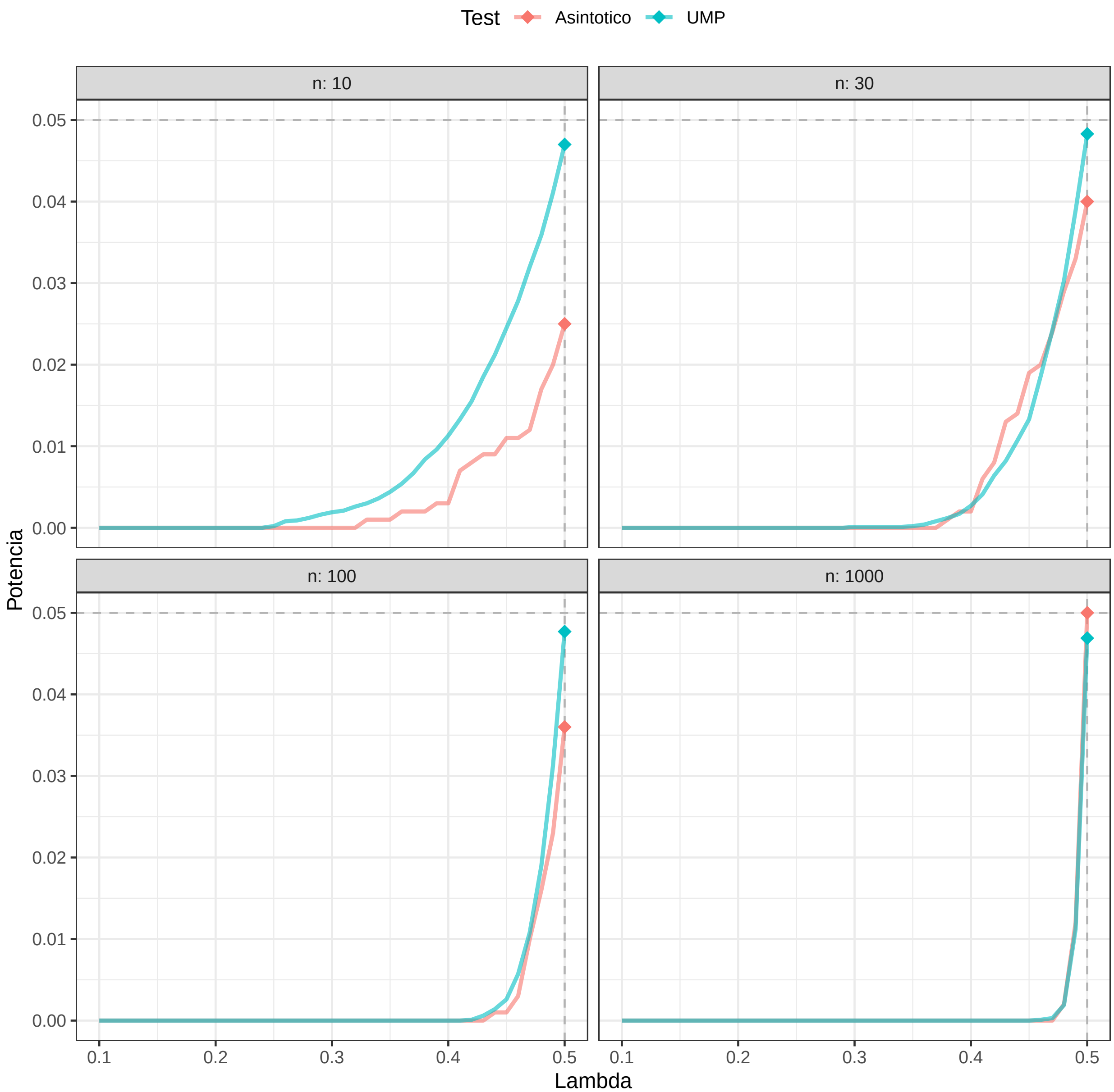
```r
# La función simulación simula el experimento y se fija el resultado del test
set.seed(123)
ns <- c(10,30,100,1000) # Posibles ns
lambdas <- seq(0.1,1,.01) # Rango de lambdas

potencias <- expand_grid(n = ns, Lambda = lambdas) %>%
  rowwise() %>%
  mutate(UMP = simulacion_UMP(Nrep = 10000, n = n, lambda = Lambda),
         Asintotico = simulacion_asin(Nrep = 1000, n = n, lambda = Lambda)) %>%
  pivot_longer(cols = !c(n, Lambda), names_to = "Test", values_to = "Potencia")
```



En la figura podemos ver la función potencia $\pi_\delta(\lambda)$ para el test UMP (celeste) y asintótico (rosa) en función del λ y separado para distintos valores de n. Se observa que para n pequeño (<100) hay una diferencia en la potencia empírica a favor del test UMP, principalmente para la zona de $\lambda > 0.5$ donde vemos claramente que la función de potencia del test exacto es mayor que la del test asintótico. En cambio, para valores de n más grandes las funciones de potencia resultan muy parecidas.

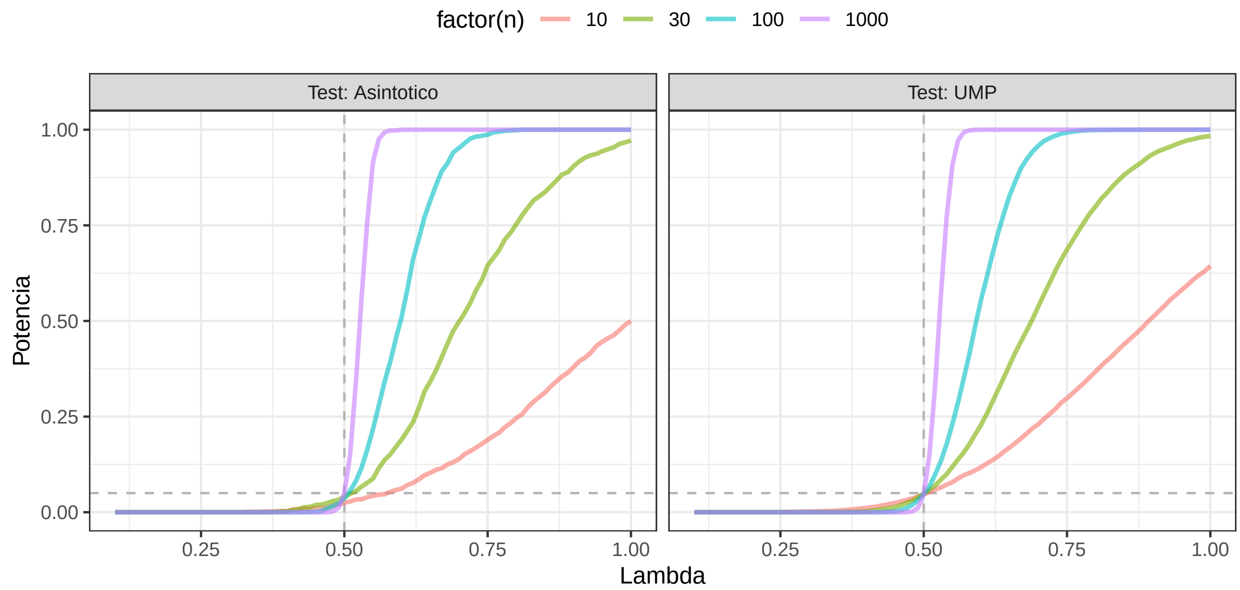
Si quisiéramos ver con mayor detalle qué pasa en la región de Θ_1, podemos hacer un zoom como se muestra a continuación.



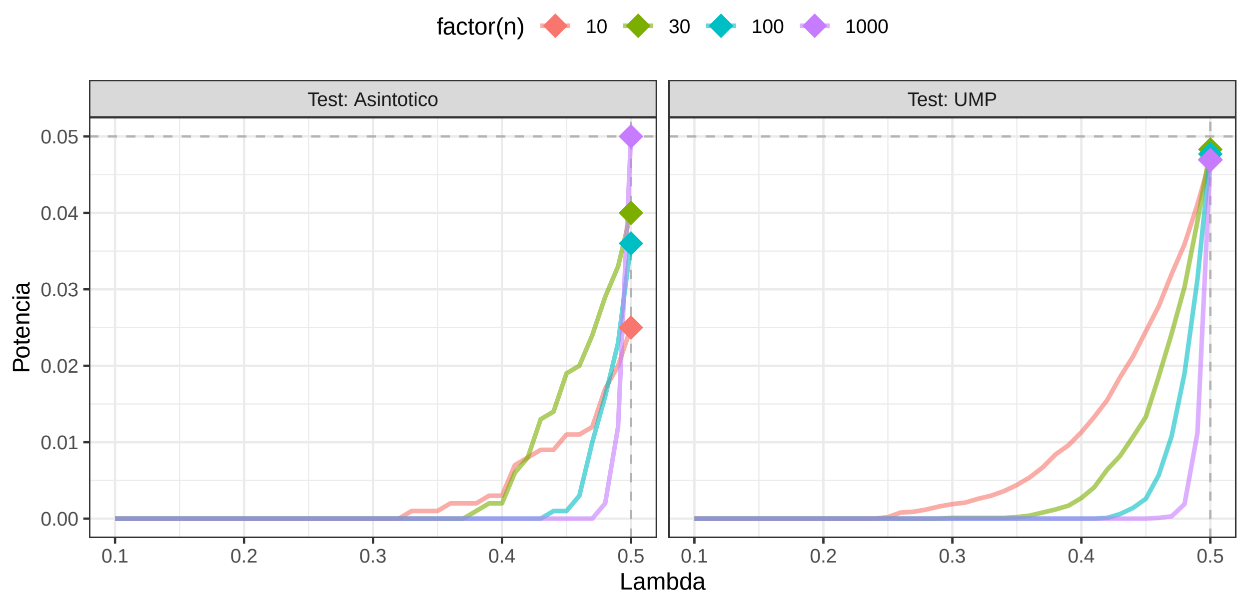
Para $\lambda < 0.5$ resulta un poco más complejo analizar qué test es "mejor". Tomemos como ejemplo $n = 10$: si bien en este caso la función de potencia parece indicar que el test asintótico es más conveniente, ya que la probabilidad de rechazo cuando nos encontramos en H_0 es más baja que la del test exacto, vemos que la potencia para $\lambda = 0.5$ no vale 0.05 como queríamos al diseñar el test, sino aproximadamente la mitad[1]. Entonces, esta mejora observada en el error de Tipo I tiene la contraparte de una menor probabilidad de rechazo cuando H_0 sí es falsa (es decir, a partir de 0.5 en λ), inflando el error de tipo II, como puede verse en la figura anterior. Es por eso que consideramos que para n bajos, el test UMP sigue siendo mejor que el asintótico. Por otro lado, al igual que en el caso de $\lambda > 0.5$, podemos ver que para $n > 1000$ ambos test son indistinguibles.

Otro aspecto que podemos analizar es cómo varía la potencia para cada test con el n.

[1]Notar que estos graficos de potencia fueron realizados con unas simulaciones diferentes a los de la tabla del inciso c), y es por eso que sus valores para $\lambda = 0.5$ difieren ligeramente (0.22 y 0.25).



En este caso la conclusión principal (y algo trivial) es que al aumentar el n la potencia "mejora", dado que la pendiente es más empinada cerca de 0.5. Y lo mismo podemos ver haciendo un zoom para los $\lambda < 0.5$, como se muestra en el siguiente gráfico.



e. Concluir en base a los resultados observados.

A partir de los resultados obtenidos podemos concluir que ambos tests garantizan que la proporción de rechazo bajo H_0 sea menor que 0.05 para los distintos valores de n estudiados. Para valores de $n < 100$ el nivel del test asintótico resulta más bajo que el del test exacto, pero observamos que el nivel del test exacto se aproxima mejor al deseado. A su vez, la potencia para valores de λ mayores que 0.5 es mejor en el caso del test exacto y, en este sentido, consideramos que podría resultar más conveniente utilizar este último. Por otra parte, para valores mayores de n, el desempeño de ambos tests resulta muy similar.