
Reglas del TP:

- Este trabajo debe hacerse de forma individual o de a dos.
- Deben enviarse el script de R y un informe que contenga las respuestas a todas las preguntas y los gráficos pedidos. No hace falta explicar en el informe que es lo que hace cada una de las funciones del script.
- El script debe estar prolijo. Esto en particular implica que las variables tienen que tener un nombre descriptivo (es decir, no llamar a, b, c a las variables).

El problema del candidato



Supongamos que somos el jefe de recursos humanos de una empresa y queremos contratar a alguien para una determinada posición. Tenemos n candidatos haciendo una fila esperando para entrar a la entrevista. Estos candidatos están rankeados del mejor al peor (digamos que el peor es 1 y el mejor es n), pero nosotros todavía no conocemos este ranking. Una vez que hacemos pasar a alguien a la entrevista, podemos conocer sólomente el ranking relativo entre este candidato y las personas que ya fueron entrevistadas. Una vez que se entrevistó a alguien, hay que darle una decisión en el momento: se lo puede contratar (y en este caso no se entrevista más a nadie) o rechazarlo y entrevistar al siguiente en la fila (el rechazado no puede volver). Nuestro objetivo es contratar al mejor entre los n candidatos. Las decisiones que tomemos sólomente pueden basarse en lo que hayamos visto hasta ese momento.

Si $n > r \geq 1$, se define la estrategia $\mathcal{E}_{n,r}$ de la siguiente forma: se entrevista a las primeras r personas de la fila y se las rechaza a todas sin importar su ranking relativo. Luego, se contrata al primer candidato que sea mejor que todos los r postulantes que se rechazaron al principio. Si ninguno de ellos es mejor que los r iniciales, entonces se contrata al último de la fila. Por ejemplo, si $n = 6$ y el ranking de los candidatos es 134256 (en el orden de la fila), entonces con la estrategia $\mathcal{E}_{6,3}$ se rechaza automáticamente a los primeros tres candidatos (con rankings 1, 3 y 4) y se contrata al quinto candidato (porque es el primero con mejor ranking que los tres candidatos iniciales de la fila). En este

ejemplo particular no se cumple el objetivo, es decir, no se contrata al mejor candidato de la fila (que era el último). Llamemos $p_{n,r}$ a la probabilidad de que se elija al mejor de los n candidatos siguiendo la estrategia $\mathcal{E}_{n,r}$.

1. Intuitivamente, ¿qué pasa con $p_{n,r}$ si r es muy chico? ¿Y si r es muy cercano a n ?
2. Hacer una función que aproxime $p_{n,r}$ mediante una simulación Monte Carlo. ¿En qué resultado teórico se basa esta aproximación?
3. Para $n = 50$, hacer un gráfico que compare r con los valores aproximados de $p_{n,r}$ ($r = 2, \dots, 49$).
4. Sea $r_n = \operatorname{argmax}_r p_{n,r}$. Hacer una función que calcule, usando las aproximaciones anteriores, el valor r_n para $n = 1, \dots, 50$. Hacer un gráfico que compare n con el valor r_n . ¿Qué tipo de dependencia nota entre las dos variables?
5. Mostrar empíricamente con los resultados del ítem anterior que

$$\lim_{n \rightarrow \infty} \frac{n}{r_n} = e. \quad (1)$$

Conjeturar cuál es el límite $\lim_{n \rightarrow \infty} p_{n,r_n}$.

6. (Opcional) Extender este trabajo en la dirección que quiera.
7. (Opcional, difícil) Probar analíticamente que vale el límite (1).
Ayuda: hallar una fórmula exacta para $p_{n,r}$.
8. (Opcional, demasiado difícil) Encontrar una estrategia mejor que \mathcal{E}_{n,r_n} (es decir, la probabilidad de elegir el mejor candidato con la nueva estrategia debe ser mayor que usando \mathcal{E}_{n,r_n}).