



# GRANDEZA EN EL FUTBOL

Enfocado desde las Simulaciones

---

*Lucas Giúdice, Juan Manuel Karawacki y Bruno Pintos*

# OBJETIVO GENERAL

ANALIZAR CÓMO SE DISTRIBUYEN LOS TÍTULOS EN EL LARGO PLAZO Y USAR ESA DISTRIBUCIÓN COMO INDICADOR DE “GRANDEZA” RELATIVA ENTRE LOS CLUBES.

EXPLORAR EL IMPACTO DE DISTINTAS VARIABLES EN UN TORNEO A LO LARGO DEL TIEMPO:

- PRESUPUESTO
- LESIONES
- PROBABILIDADES DE GOL
- ESTRUCTURA DE PREMIOS





# PREGUNTAS PLANTEADAS

**¿Cómo se reparten los campeonatos entre clubes cuando combinamos factores futbolísticos, económicos y azar?**

**¿Cuanto cuesta ganar un punto mas en la liga segun el presupuesto?**

**¿Como se distribuye el costo por punto?**

**¿Qué cambia entre:**

- una liga donde todos arrancan con presupuestos similares y**
- una liga consolidada donde el presupuesto se concentra en pocos equipos?**

# ESTRUCTURA GENERAL DE LA LIGA SIMULADA: EQUIPOS

- 20 equipos.
- Cada equipo cuenta con 22 jugadores: 11 titulares y 11 suplentes.
- Se dividen en 4 posiciones ocupadas en el campo: goleiros, defensas, mediocampistas y delanteros.





# ESTRUCTURA GENERAL DE LA LIGA SIMULADA: DINÁMICA

- 38 partidos cada equipo; enfrentándose 2 veces entre sí.
- Premiación económica decreciente para los 10 primeros equipos del torneo. Considerados en temporadas futuras.
- Sistema de competencia tradicional: 3 puntos por victoria, 1 por empate, 0 por derrota.
- Desempate final por goles a favor/diferencia de goles.

# COMPONENTES DE LOS DATOS

→ A nivel equipos:

- Presupuestos anuales.
- Dimensión del club.

→ A nivel jugadores:

- Probabilidades de gol.
- Probabilidades de lesión.



# EQUIPOS: PRESUPUESTO Y DIMENSIÓN

**Idea:** Modelar una realidad de presupuestos dispares.

La dimensión de los equipos determina su “*intervalo presupuestal*”:

- **Equipos “grandes”:** entre 800 y 1.200.
- **Equipos “chicos”:** entre 20 y 400.

Permitir flexibilidad con propósitos experimentales.



# JUGADORES: PROBABILIDAD DE LESIÓN

A la hora de asignar una probabilidad de lesión se toma en cuenta:

- Posición de jugador.
- Grandeza del equipo → Ajuste por presupuesto (cantidad de partidos).
- Titularidad.

La asignación es arbitraria.



# ¿Y EL MÉTODO NUMÉRICO?

# ¿Y EL MÉTODO NUMÉRICO?



# ¿Y EL MÉTODO NUMÉRICO?



**EXACTO!! : Probabilidades de gol**

# JUGADORES: PROBABILIDAD DE GOL

A la hora de asignar una probabilidad de gol se toma en cuenta:

- Posición de jugador.
- Grandeza del equipo.
- Titularidad

Si bien la asignación es arbitraria, requerimos de una posterior calibración:

→ **Ajustamos con datos reales**

# ¿POR QUÉ NECESITAMOS CALIBRAR EL MODELO?

- El objetivo es reproducir un promedio real de 2.61 goles por partido.
- Las probabilidades de gol por posición (GK, DF, MF, FW) provienen de datos reales, pero necesitan un ajuste global.
- $\theta$  es un parámetro global de escala
- El parámetro  $\theta$  controla la “intensidad goleadora” total del modelo.

Buscamos el valor de  $\theta$  tal que:

$$g(\theta) = E[(\text{sim\_lig})^*(\theta)] - 2.61 = 0$$



# **PRIMER ENFOQUE SIMULACIÓN: MONTE CARLO DIRECTA**

**Se parte de un vector de probabilidades “semilla”**

**En cada iteración:**

1. Se simulan cientos de partidos usando el motor completo (lesiones, suplentes, presupuesto).
2. Se calcula el promedio de goles observado.
3. Se ajusta  $\theta$  multiplicando por un factor correctivo.

**Itera hasta que la diferencia entre promedio simulado y observado sea pequeña.**



- Captura toda la complejidad del modelo real.
- Ajuste flexible: Cualquier dinámica interna afecta el resultado final.
- Fácil de implementar conceptualmente



## VENTAJAS DEL ENFOQUE MONTE CARLO

# RESULTADO DE LA CALIBRACIÓN MONTE CARLO

- Tras varias iteraciones, se obtiene un  $\theta$  calibrado (~0.65).
- Fácil de implementar conceptualmente.
- Sirve como estimación inicial para el método numérico posterior



# COMPARACIÓN: MÉTODO DE BISECCIÓN

Buscamos el valor de  $\theta$  tal que:  $g(\theta) = E[G(\theta)] - 2.61 = 0$

Requiere únicamente:

1. Un intervalo  $[a,b]$  tal que  $g(a)g(b) < 0$ .
2. Poder evaluar  $g(\theta)$ .

Avanza dividiendo sucesivamente el intervalo hasta aproximar la raíz.



# ¿POR QUÉ BISECCIÓN Y NO NEWTON O PUNTO FIJO?



# RESULTADO CON BISECCIÓN

- En ~7 iteraciones se obtiene:

**Valor aproximado:**  $\theta^* \approx 0.653$

**Diferencia residual:**  $g(\theta^*) \approx 0.003$



# INICIAMOS LA SIMULACIÓN



# SIMULACIÓN: PARTIDOS

En cada partido sucede:

- Se calcula el **ratio entre el presupuesto de aquel que más posee y el que menos**. Eso se utilizará para torcer las probabilidades de gol en favor del más rico.
- Se **simulan eventos de lesión** para los jugadores sanos. Los **suplentes no pueden lesionarse** (esto nos asegura que siempre tenemos jugadores disponibles)
- Si **faltan titulares se reemplazan** con suplentes de dicha posición.
- Se calculan **eventos de gol** a nivel de los 11 jugadores de cada equipo. Aquel cuyos jugadores hayan metido más goles en total **gana el encuentro**. Se resaltan eventos atípicos.

# TRANSCURSO DE LA LESIONES

## Probabilidades base por posición:

- GK: 0.0025
- DF: 0.0075
- MF: 0.01
- FW: 0.01

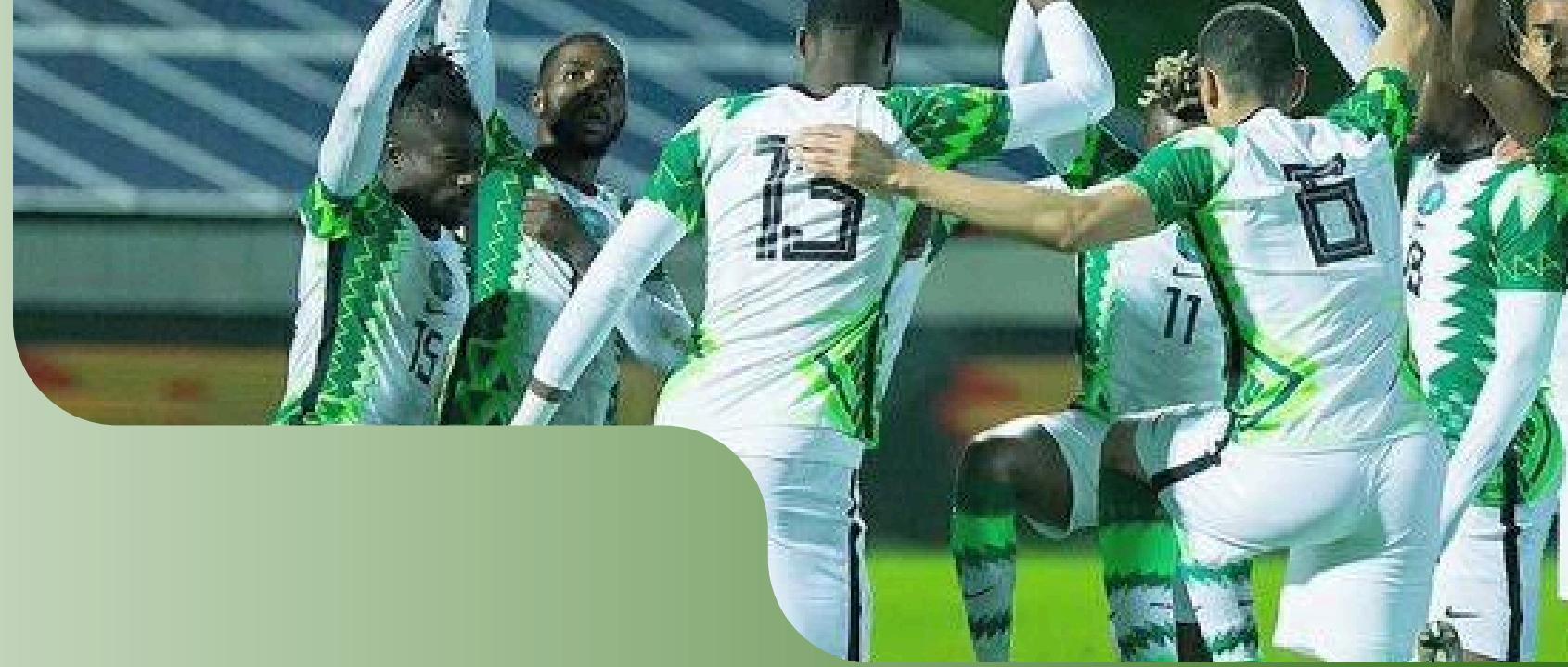
## Dinámica por partido:

- Se simulan nuevas lesiones solo en jugadores sanos.  
Probabilidad Bernuolli.
- Duración de nuevas lesiones: aleatoria entre 1 y 8 partidos.
- Partido a partido se reduce la duración de lesiones existentes (-1 partido hasta 0).

# SIMULACIÓN: LIGA

En cada liga / temporada sucede:

- Se **genera un fixture** de encuentros visitante local para todos los equipos.
- Se **itera sobre dicho fixture**, recorriendo cada fecha para los equipos. En cada una se simulan partidos.
- Se lleva un **recuento de los jugadores que se van lesionando** y de los procesos de recuperación de los que ya lo están.
- Por fecha se juegan 10 partidos, con la posibilidad de ver los resultados específicos de cada uno para cada fecha.
- Al final del torneo **triunfa el equipo con más puntos**, y más goles en caso de empate



# GRANDEZA



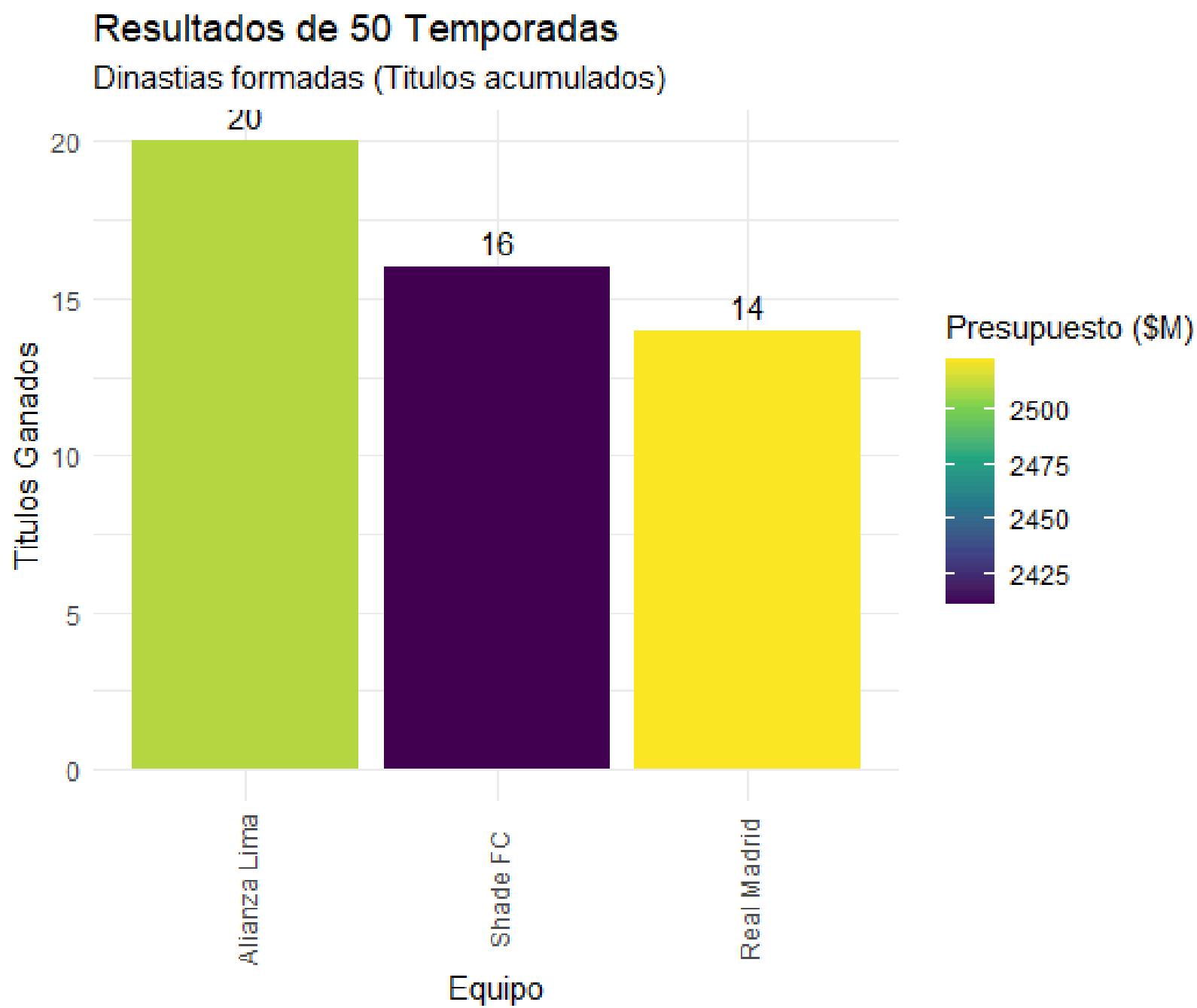
Para visualizar la grandeza de los clubes, queremos ver **como sostienen sus triunfos a lo largo del tiempo:**

→ **Simulamos n ligas jugadas con acumulación una sobre otra.**

Esto implica que la estructura de premios concedidos en cada temporada, influye en los rendimientos de cada una de las ligas futuras.

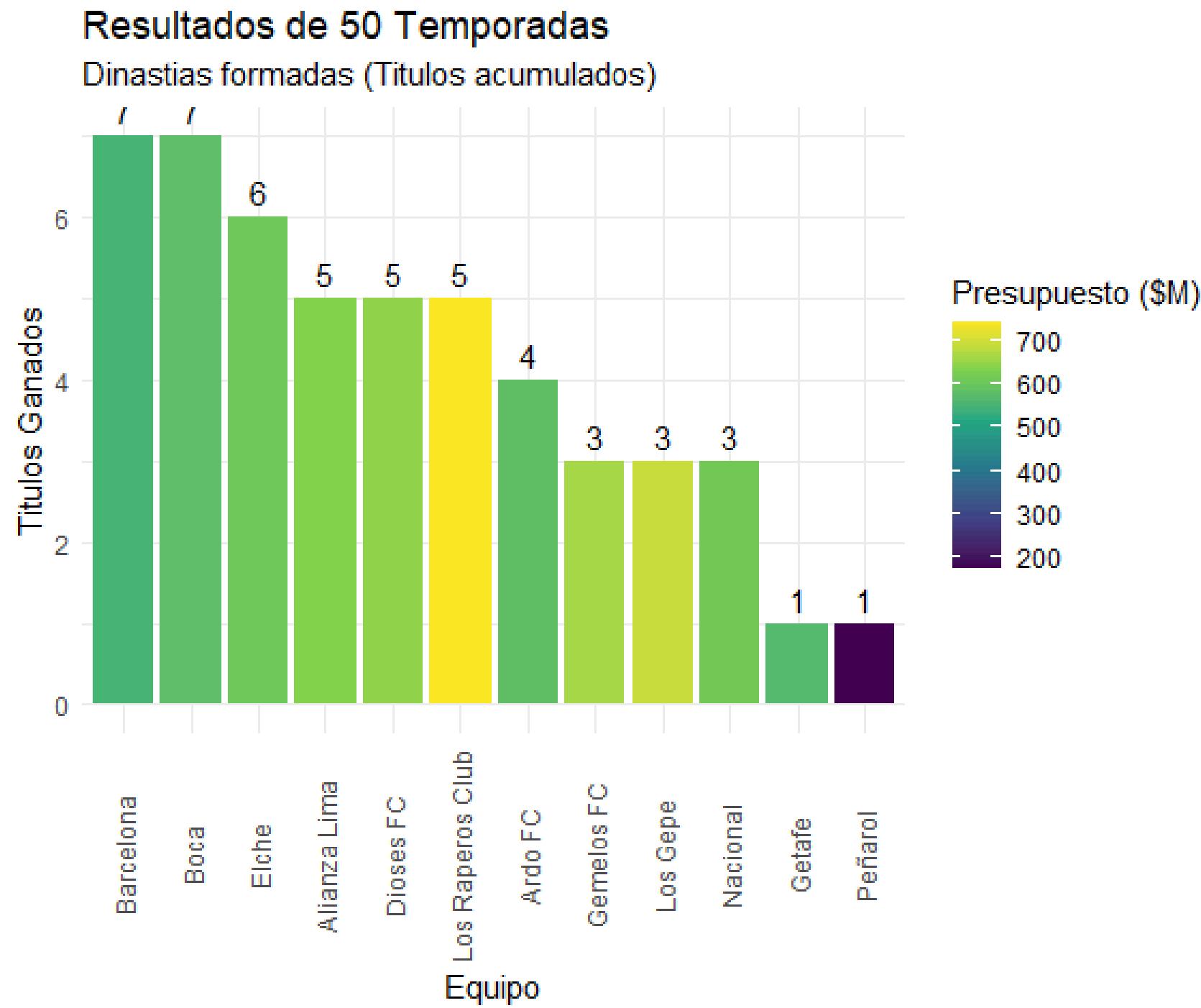
# RESULTADOS

# 50 TORNEOS: TRES EQUIPOS GRANDES



Cuando tenemos pocos equipos que concentran aproximadamente el 80% del presupuesto en la liga, vemos que son los únicos que terminan ganando los títulos al final de la liga.

# 50 TORNEOS: SIN EQUIPOS GRANDES



Cuando todos los equipos comienzan con los mismos presupuestos, los títulos de liga se reparten mucho mas y vemos el crecimiento presupuestal.

# OPTIMIZACIÓN

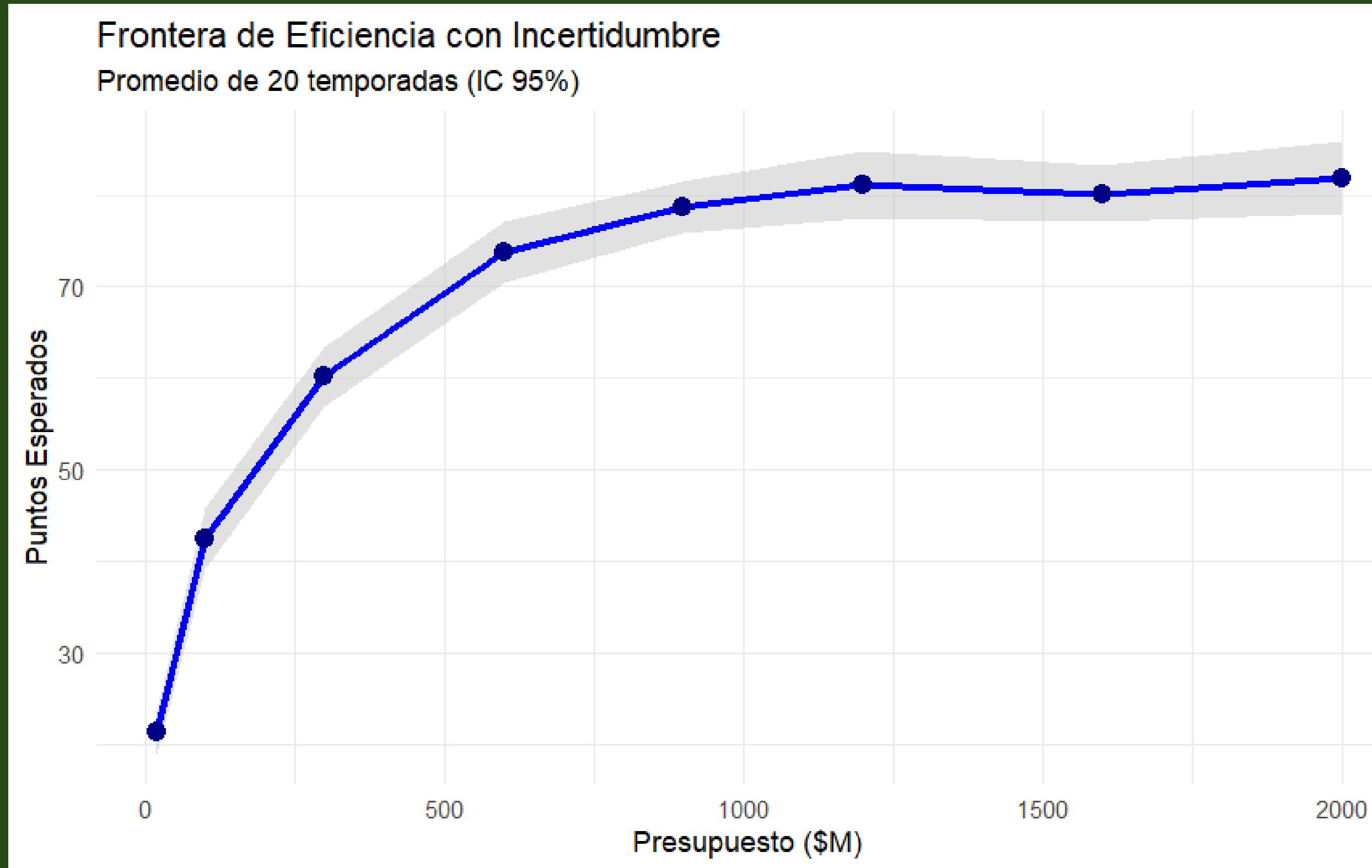
Si duplicamos el presupuesto de nuestro equipo hoy mismo, ¿garantizamos duplicar nuestros puntos?

¿Cuánto cuesta comprar “un punto más” en la tabla de posiciones?

¿En qué punto exacto invertir más dinero se vuelve una mala decisión?



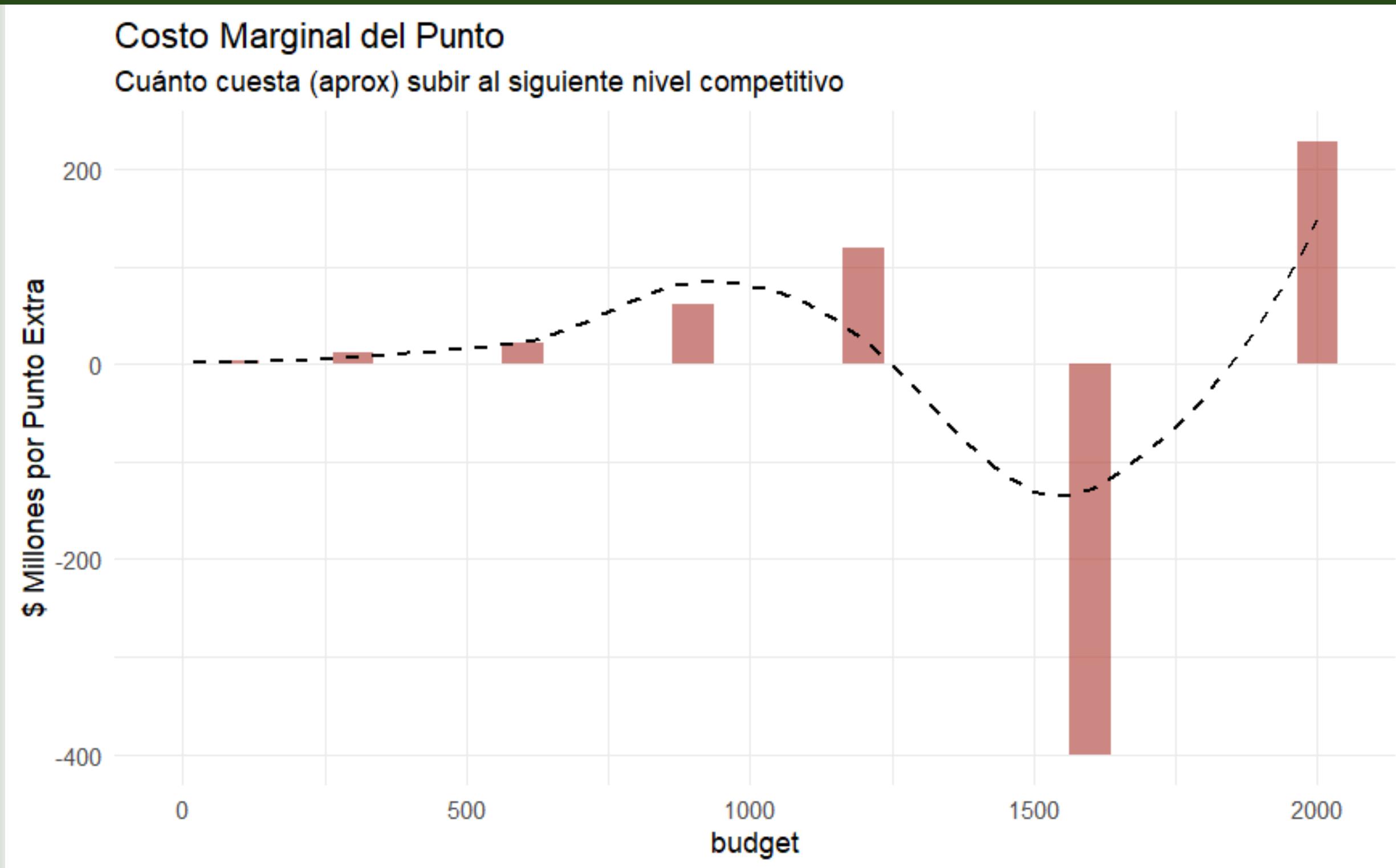
# RELACION: DINERO - PUNTOS



Al principio cada punto adicional es baratisimo

Luego, a pesar de aumentar muchos millones los puntos se estancan...

# ¿CUANTO VALE CADA PUNTO?



Al principio es barato conseguir un punto extra.

A partir de los 900 millones cada punto se vuelve carísimo.

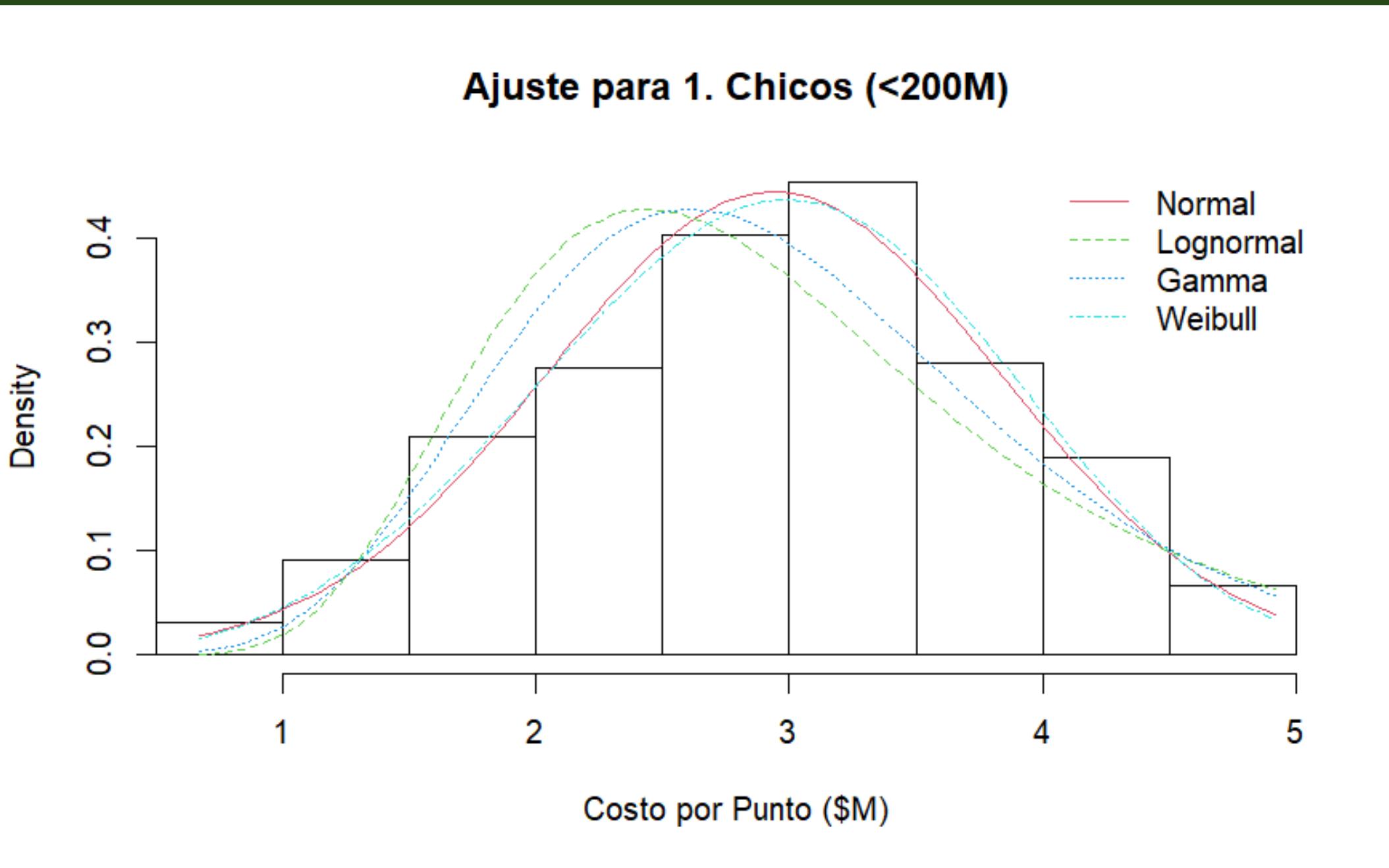
Al rededor de los 1500 incluso estamos pagando mas para obtener menos puntos que antes.

# ¿PERO COMO DISTRIBUYE EL COSTO POR PUNTO?

Ya vimos como es el promedio pero, ¿cuál es el comportamiento fuera de la media?

La librería **fitdistrplus** responde a esta duda para nuestros diferentes niveles de **presupuesto**

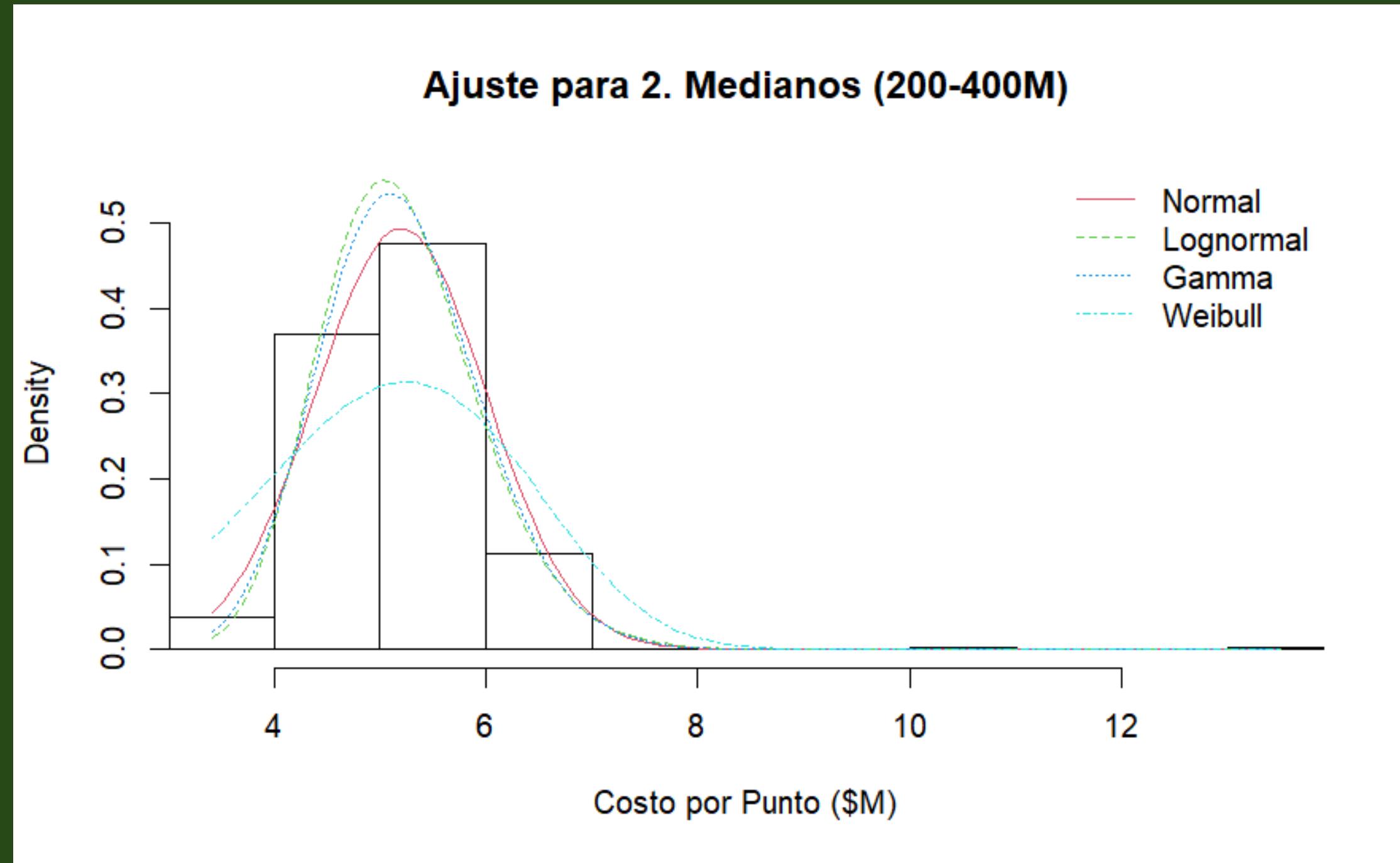
# PARA LOS EQUIPOS CHICOS



- Los costos por punto en los equipos chicos están concentrados entre 1 y 5 millones, con mayor frecuencia al rededor del 3 con cierta asimetría.

El AIC nos dice que la distribución **Weibull** es la que mejor representa el Costo por Punto de los equipos chicos.

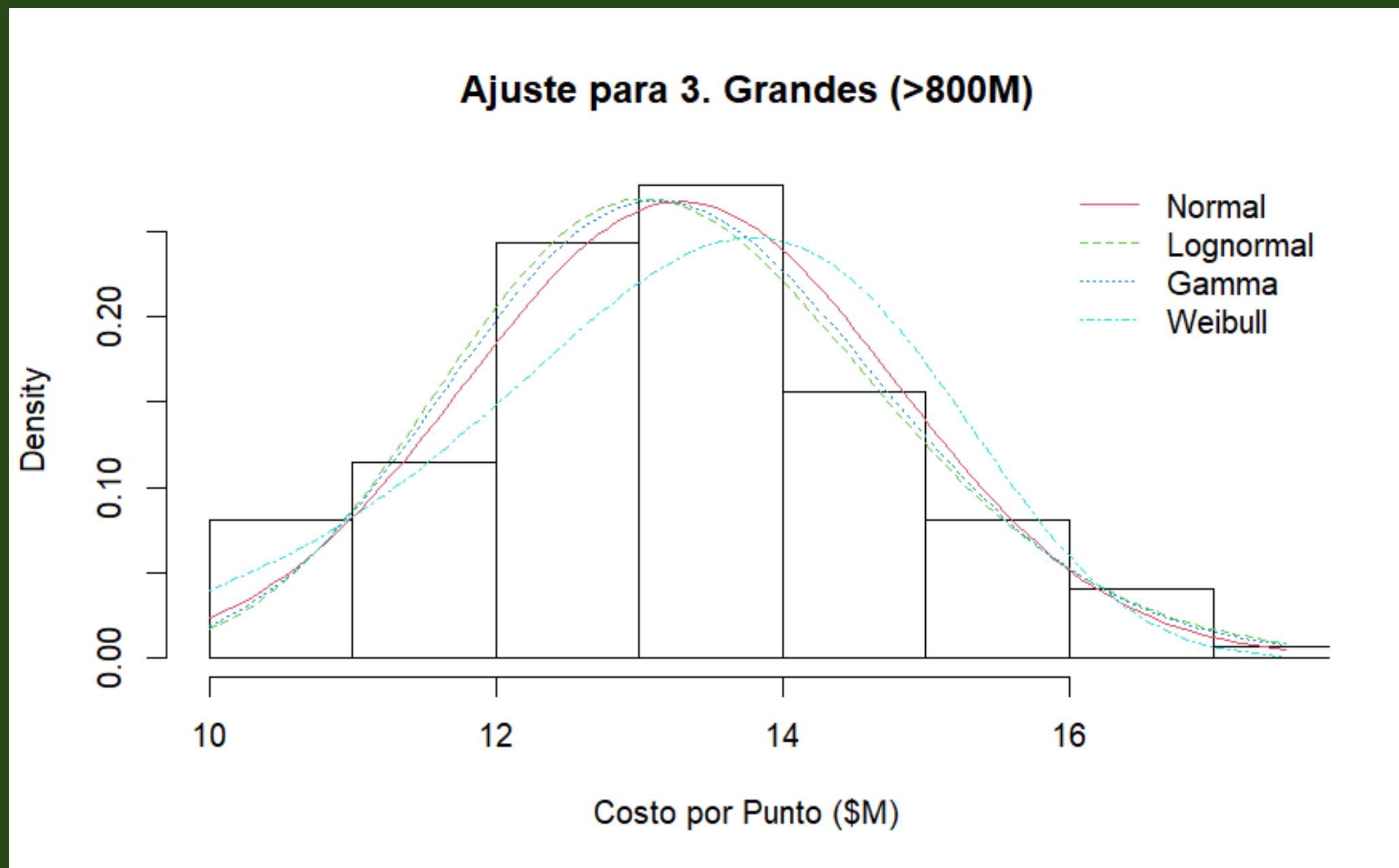
# PARA LOS EQUIPOS MEDIANOS



El AIC nos dice que la distribución **Lognormal** es la que mejor representa el Costo por Punto de los equipos medianos.

- El costo promedio sube prácticamente duplicando los costos de los equipos de presupuestos mas bajos y aparece la asimetría con una cola hacia la derecha.
- Comienzan a verse lesiones y notarse los costos marginales crecientes.

# PARA LOS EQUIPOS GRANDES



El AIC nos dice que la distribución Normal es la que mejor representa el Costo por Punto de los equipos grandes.

- Mucha mayor varianza, equipos con diferencias de 6 millones de costo \* 80 puntos en promedio: [800; 1280]
- Esto es que dos equipos con el mismo presupuesto pueden tener resultados financieros muy distintos solo por la suerte (lesiones, racha de goles)

# CONCLUSION

## 1. EQUIPOS CHICOS (<200M): LA EFICIENCIA DE SUPERVIVENCIA

El costo se concentra en un rango de máxima eficiencia (~3M/punto), al estar protegidos de la saturación (lesiones) su desempeño refleja una maximización forzada de recursos donde el despilfarro es sistémicamente imposible.

# CONCLUSION

## **2. EQUIPOS MEDIANOS (200M - 800M): LA FASE DE TRANSICIÓN Y COSTOS CRECIENTES**

En el segmento medio, la distribución se desplaza hacia la derecha. El incremento en el costo medio (~5.5M/punto) y la aparición de colas hacia costos más altos evidencian el impacto de los costos marginales crecientes. Esta fase marca la pérdida de ventajas competitivas iniciales, donde ganar puntos adicionales requiere un esfuerzo financiero desproporcionalmente mayor al de la etapa anterior.

# CONCLUSION

## 3. EQUIPOS GRANDES (>800M): SATURACIÓN Y GESTIÓN DE RIESGO

Para los equipos grandes el costo medio se estabiliza en un nivel alto estructural (~13.5M/punto) debido a los rendimientos decrecientes, y la alta varianza observada refleja que, a estos niveles de inversión, el éxito financiero depende más de la mitigación de riesgos (azar/lesiones) que de la estructura de costos base.

# MUCHAS -GRACIAS-

