**Metodo Bayesiano**

**Paso a paso**

**Definición del objetivo y la métrica**

El objetivo del experimento es predecir el horario más efectivo (AM, MD, PM) para contactar a un cliente, maximizando la probabilidad de que conteste la llamada.  
La métrica principal será la tasa de contacto, definida como:

p=nuˊmero de clientes que contestannuˊmero total de llamadas realizadasp = \frac{\text{número de clientes que contestan}}{\text{número total de llamadas realizadas}}p=nuˊmero total de llamadas realizadasnuˊmero de clientes que contestan​

Cada observación corresponde a un cliente (lead) y la respuesta es binaria: 1 si contesta, 0 en caso contrario.

**2. Planteamiento en el marco bayesiano**

En lugar de hipótesis nula y alternativa, el análisis se centrará en estimar la probabilidad posterior de que el grupo de tratamiento supere al control.  
Se calcularán:

* Pr⁡(pB>pA)\Pr(p\_B > p\_A)Pr(pB​>pA​), es decir, la probabilidad de que el tratamiento supere al control.
* Pr⁡(pB−pA>δ)\Pr(p\_B - p\_A > \delta)Pr(pB​−pA​>δ), probabilidad de que la diferencia entre ambos supere un mínimo efecto detectable (δ\deltaδ).

**3. Umbrales de decisión**

Se establecen reglas de decisión basadas en probabilidades posteriores:

* Superioridad: se considera que el tratamiento es mejor si Pr⁡(pB>pA)≥0.95\Pr(p\_B > p\_A) \geq 0.95Pr(pB​>pA​)≥0.95.
* Relevancia práctica: además, se requiere que Pr⁡(pB−pA>δ)≥0.90\Pr(p\_B - p\_A > \delta) \geq 0.90Pr(pB​−pA​>δ)≥0.90, donde δ\deltaδ representa el mínimo efecto detectable definido previamente.
* Futilidad: si la probabilidad de que el tratamiento sea mejor cae por debajo de 0.20, se interrumpe el experimento.

**4. Tamaño de muestra**

En el enfoque bayesiano no es necesario calcular un tamaño de muestra exacto mediante fórmulas cerradas. En su lugar se definen:

* Un número mínimo de observaciones por grupo (nminn\_{min}nmin​) para garantizar que la inferencia posterior no dependa únicamente de los priors.
* Un número máximo (NmaxN\_{max}Nmax​) que limita la duración y el costo del experimento.  
  Entre estos valores, el análisis se actualiza continuamente y se toman decisiones cuando los umbrales de probabilidad se cumplen.

**5. Diseño experimental**

Se asigna aleatoriamente a cada cliente a uno de los dos grupos:

* **Control (A):** horario tradicional de llamadas.
* **Tratamiento (B):** horario recomendado por el modelo de predicción.

Cada cliente se mantiene en el grupo asignado (randomización) y solo se contabiliza el primer intento de contacto para evitar sesgos.

**6. Ejecución y recolección de datos**

Durante la ejecución se registran para cada cliente: identificador, grupo asignado, horario de llamada, resultado binario (contestó o no contestó).  
Se realizan revisiones periódicas del balance entre grupos y de la calidad del registro de datos.

**7. Análisis de resultados**

Se utiliza un modelo Beta-Binomial para cada grupo.  
Si en el grupo kkk se observan xkx\_kxk​ éxitos en nkn\_knk​ llamadas, la posterior para la tasa de contacto se define como:

pk∣D∼Beta(α+xk,β+nk−xk)p\_k \mid D \sim \text{Beta}(\alpha + x\_k, \beta + n\_k - x\_k)pk​∣D∼Beta(α+xk​,β+nk​−xk​)

donde α\alphaα y β\betaβ son los parámetros del prior (usualmente α=β=1\alpha=\beta=1α=β=1, prior no informativo).

Se generan simulaciones Monte Carlo de pAp\_ApA​ y pBp\_BpB​ para estimar:

* La probabilidad de que el tratamiento sea superior al control.
* La distribución de la diferencia Δ=pB−pA\Delta = p\_B - p\_AΔ=pB​−pA​.

**8. Intervalos de credibilidad**

A partir de la distribución posterior de Δ\DeltaΔ, se calcula el intervalo de credibilidad del 95% usando los percentiles 2.5 y 97.5.  
Este intervalo se interpreta como el rango dentro del cual se encuentra la verdadera diferencia entre tratamiento y control con un 95% de probabilidad.

**9. Decisión**

La decisión se toma aplicando las reglas predefinidas:

* Si Pr⁡(pB>pA)≥0.95\Pr(p\_B > p\_A) \geq 0.95Pr(pB​>pA​)≥0.95 y Pr⁡(Δ>δ)≥0.90\Pr(\Delta > \delta) \geq 0.90Pr(Δ>δ)≥0.90, se concluye que el tratamiento es mejor y se implementa el modelo para definir el horario de llamadas.
* Si la probabilidad de superioridad cae por debajo de 0.20, se considera que el modelo no aporta valor y se detiene el experimento.
* En caso contrario, se continúa hasta alcanzar NmaxN\_{max}Nmax​.

Aparte

La **principal diferencia** entre estos dos tipos de pruebas es que las pruebas de dos colas pueden mostrar evidencia de que el control y la variación son diferentes, mientras que las pruebas de una cola pueden mostrar evidencia de que la variación es mejor que el control.

¿Por qué elegirías uno sobre otro? Como se mencionó anteriormente, **la prueba bilateral muestra evidencia de que el control y la variación son *diferentes* , pero la prueba unilateral muestra evidencia de que la variación es *mejor que* el control.**

<https://www.canva.com/design/DAGwci1oIxI/O9dzJMQNoOnJA6n2SqUy4g/edit?utm_content=DAGwci1oIxI&utm_campaign=designshare&utm_medium=link2&utm_source=sharebutton>