

Distribuciones y parámetros y distribuciones de parámetros

Dr. Héctor Nájera
Dr. Curtis Huffman

La clase anterior

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, 142(2), 573.

Journal of Experimental Psychology: General
2013, Vol. 142, No. 2, 573–603

© 2012 American Psychological Association
0096-3445/13/\$12.00 DOI: 10.1037/a0029146

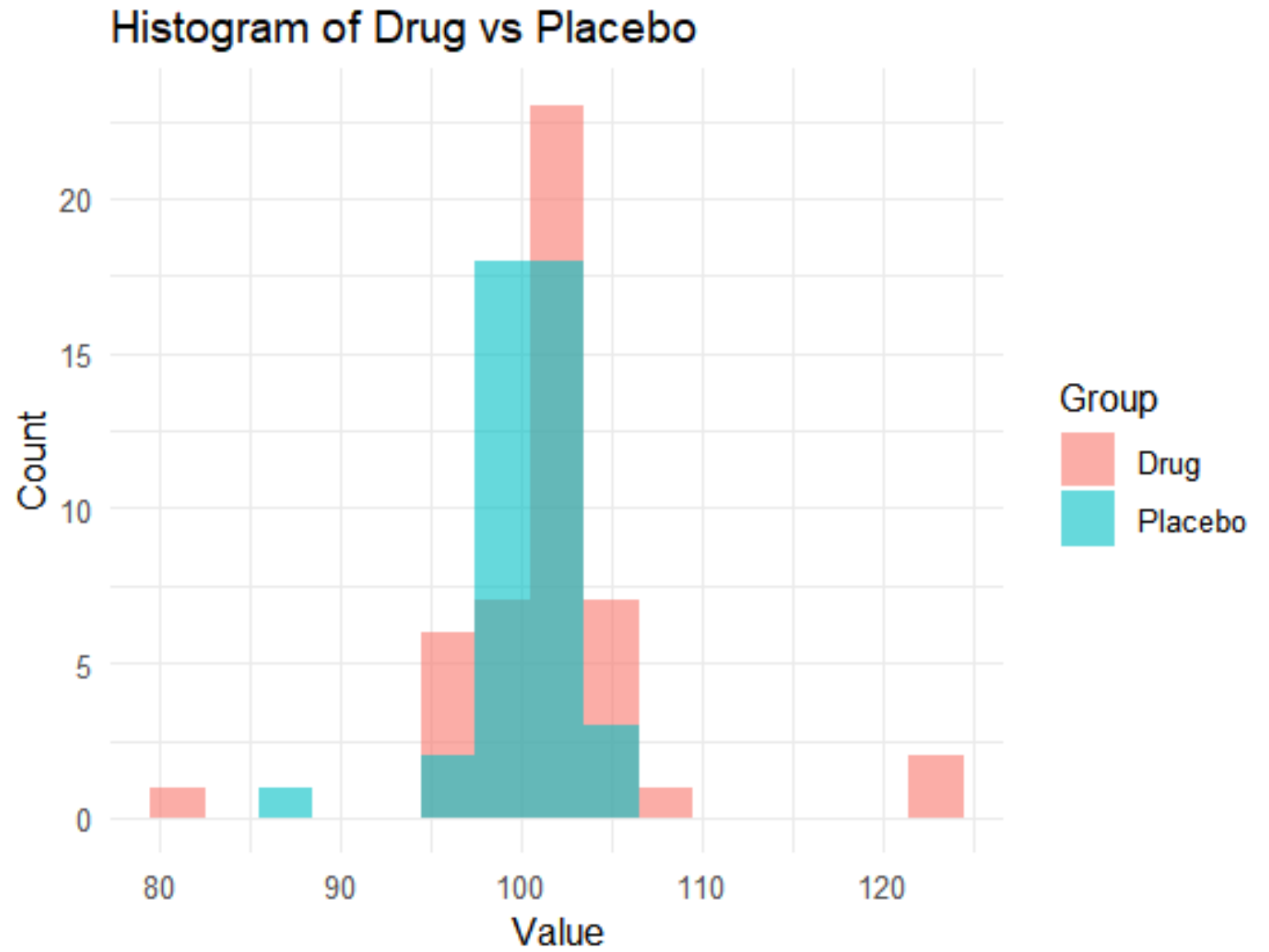
Bayesian Estimation Supersedes the t Test

John K. Kruschke
Indiana University, Bloomington

Bayesian estimation for 2 groups provides complete distributions of credible values for the effect size, group means and their difference, standard deviations and their difference, and the normality of the data. The method handles outliers. The decision rule can accept the null value (unlike traditional t tests) when certainty in the estimate is high (unlike Bayesian model comparison using Bayes factors). The method also yields precise estimates of statistical power for various research goals. The software and programs are free and run on Macintosh, Windows, and Linux platforms.

Keywords: Bayesian statistics, effect size, robust estimation, Bayes factor, confidence interval

Me interesa **generalizar** el tamaño del efecto de cierta droga sobre la inteligencia a partir de una muestra de datos de dos grupos.

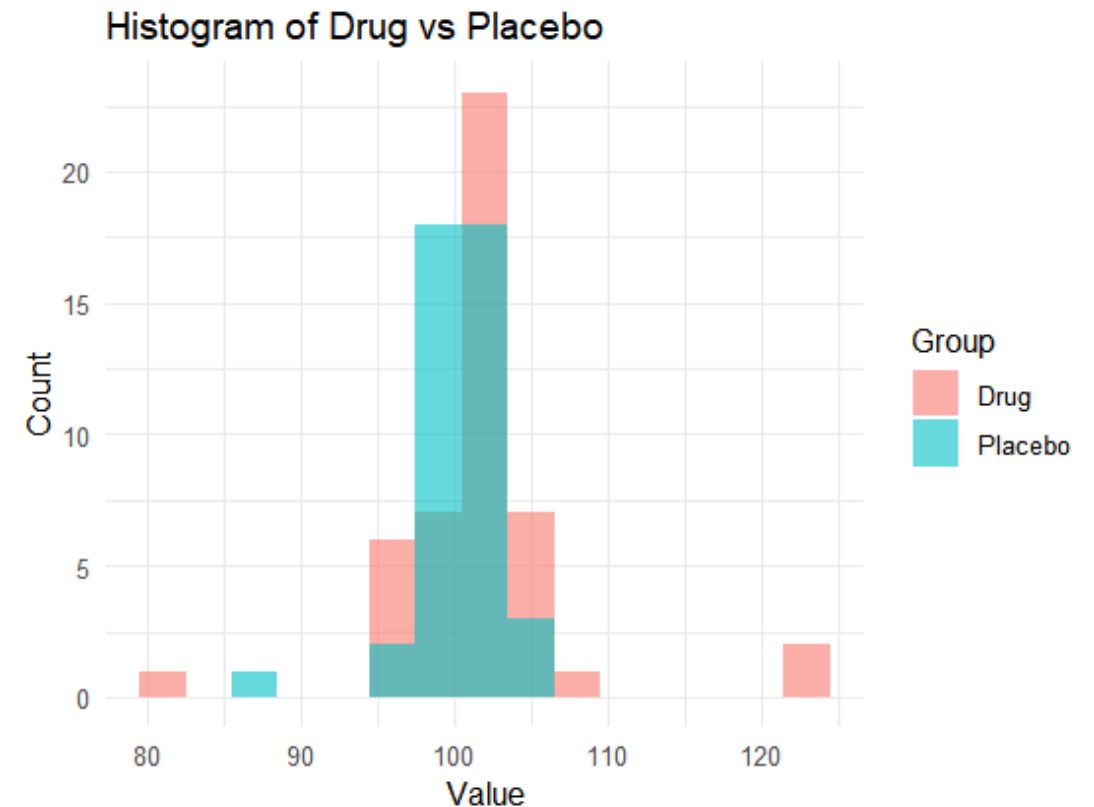


Generalizar

- Medidas explícitas y fáciles de interpretar de **qué tanto debemos de creer (incertidumbre)**, dados los datos, que cierta hipótesis (la nula u otra, no importa) se sostiene.
- Medidas que emerjan en congruencia de un marco de inferencia probabilístico que atienda estos desafíos

Modelos, distribuciones y generalizaciones

¿Por qué no basta con comparar **visual o aritméticamente** las distribuciones de los datos observados para poder **generalizar** mis hallazgos respecto a H?



Desarrollo de modelo para la generalización

- Una representación sintética y descriptiva de los datos
- Los parámetros:
 - Momento 1: son la gramática de los modelos y codifican nuestras hipótesis
 - sus distribuciones nos permiten reubicar la credibilidad en función de los datos
 - Momento 2: tienen una métrica para hacer conclusiones explícitas en función de la magnitud de sus valores probables
- Los modelos estadísticos ponen en juego estas relaciones entre las distribuciones de los parámetros y los datos (Del momento 1 al 2).



Uso de modelos estadísticos para generalizar



Los modelos estadísticos, por sí solos, no tienen alguna intención (buena o mala) por definición

Son simples ejecutores de instrucciones que les damos

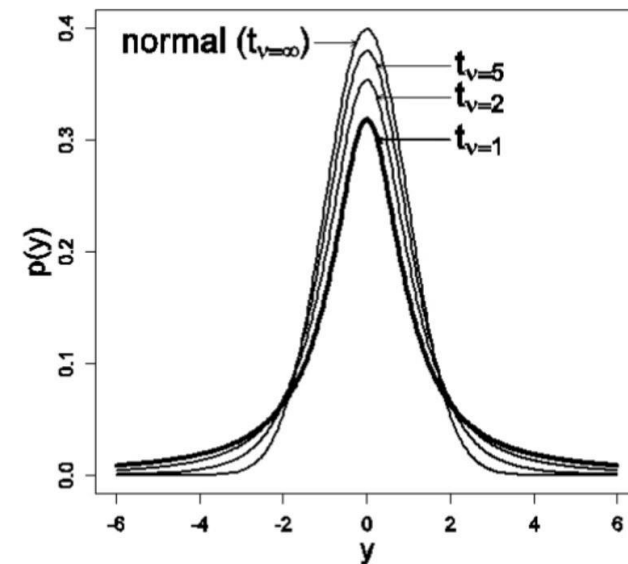
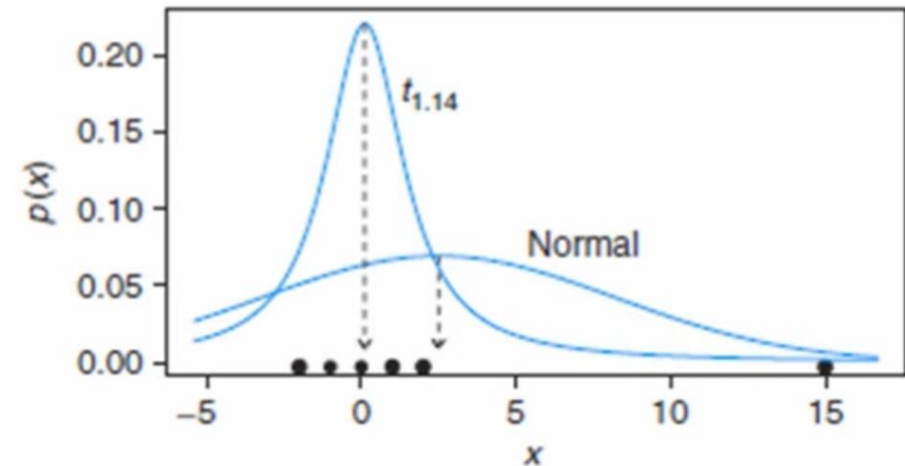
Son ciegos a nuestras muy buenas intenciones

Pueden hacer daños muy severos

¿Cómo realizar control de daños?
Modelo generativo

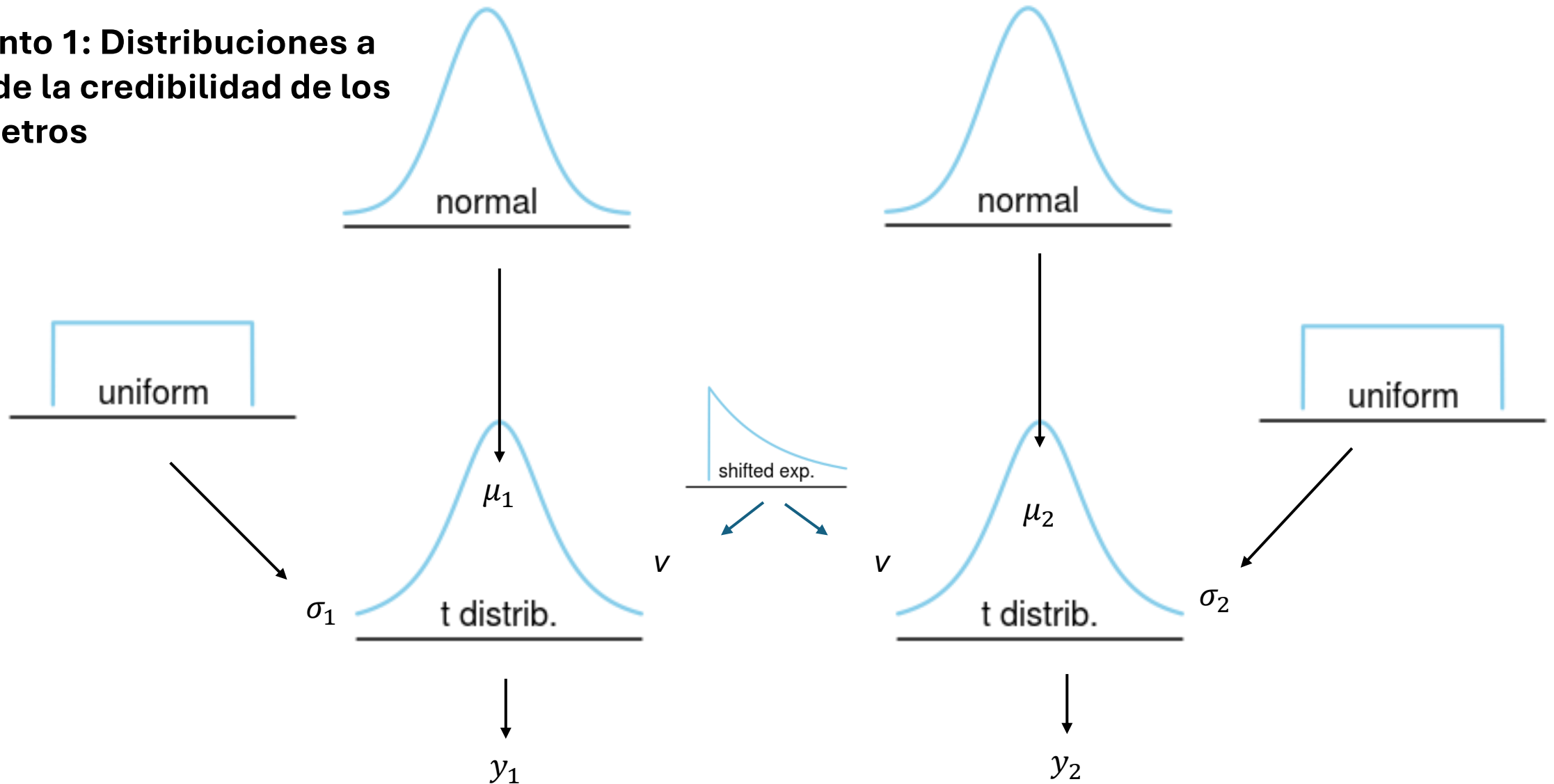
Modelo descriptivo de los datos

- La distribución normal no representa bien datos con valores extremos
- La distribución t se usa aquí para describir los datos, no como una distribución muestral de p-values



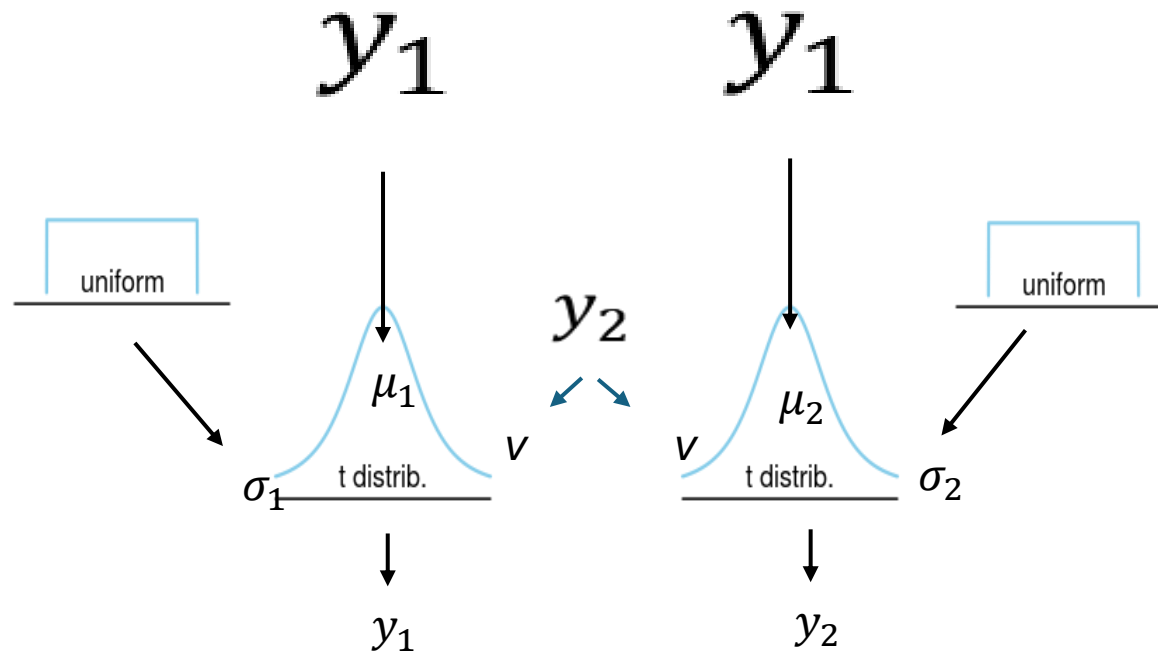
Modelo estadístico generativo de cinco parámetros

Momento 1: Distribuciones a priori de la credibilidad de los parámetros

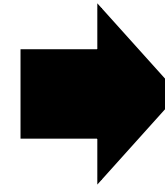


Modelo estadístico-generativo

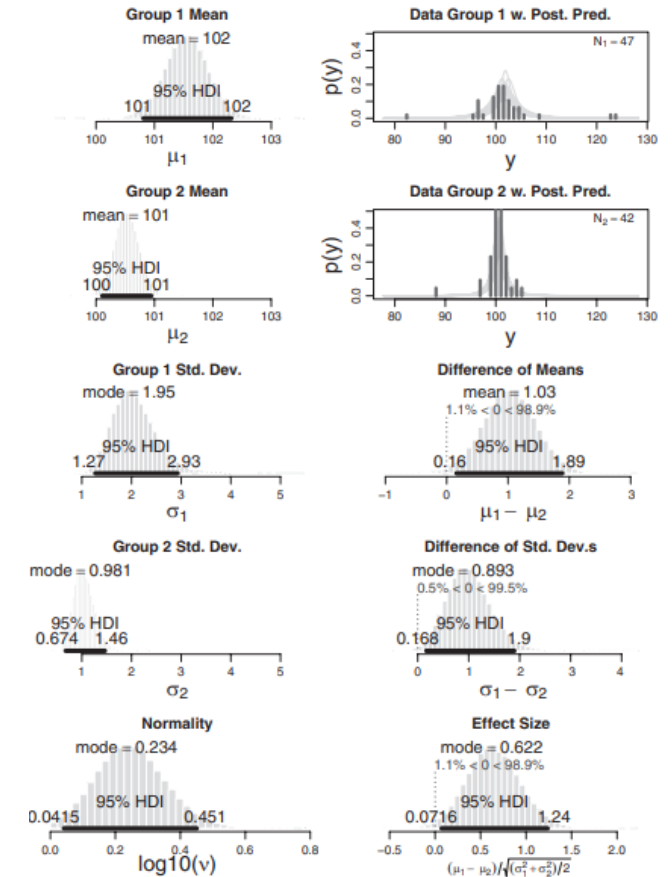
Momento 1:



¿Reubicación de credibilidad?

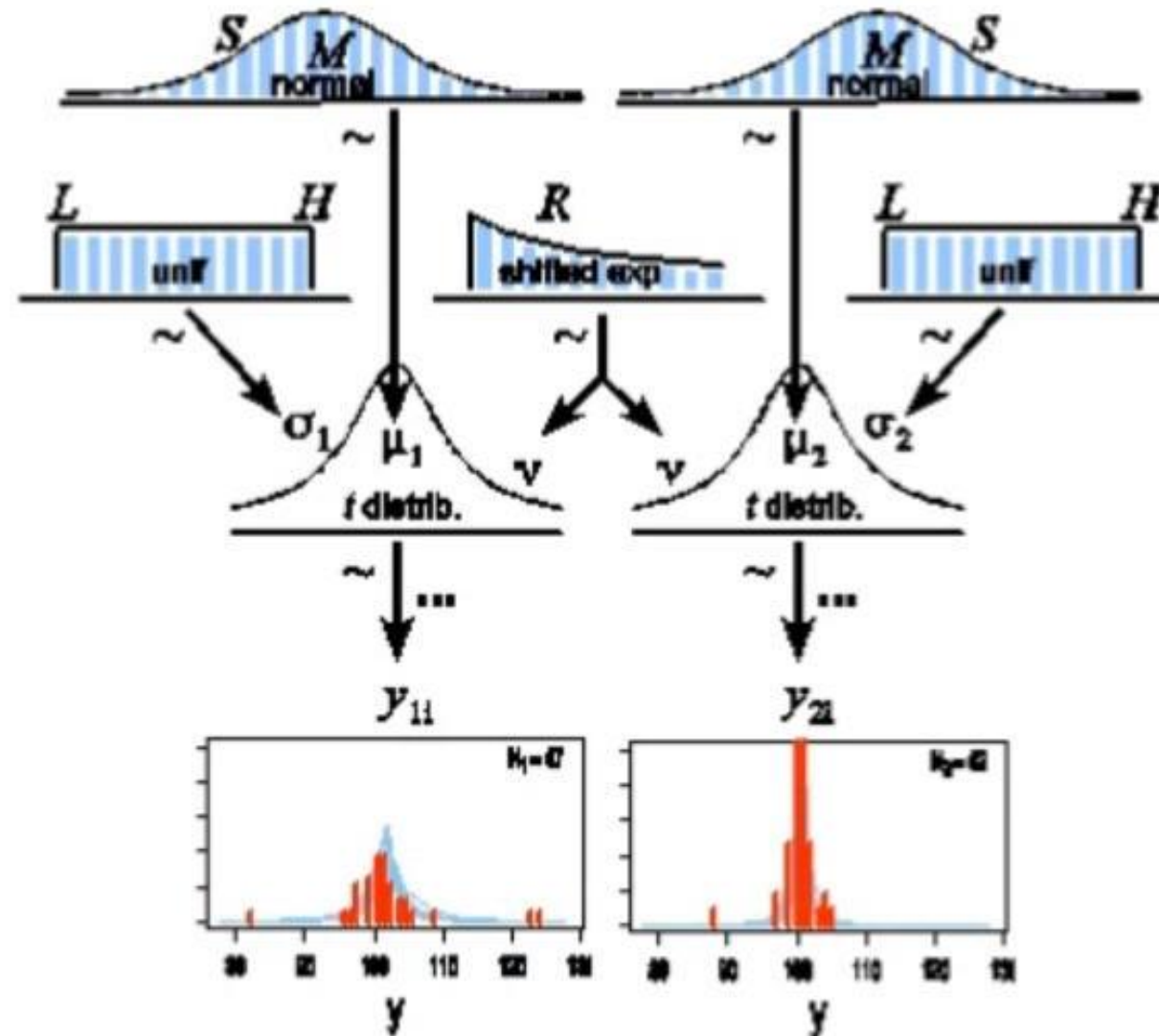


$$P(\Theta | D)$$



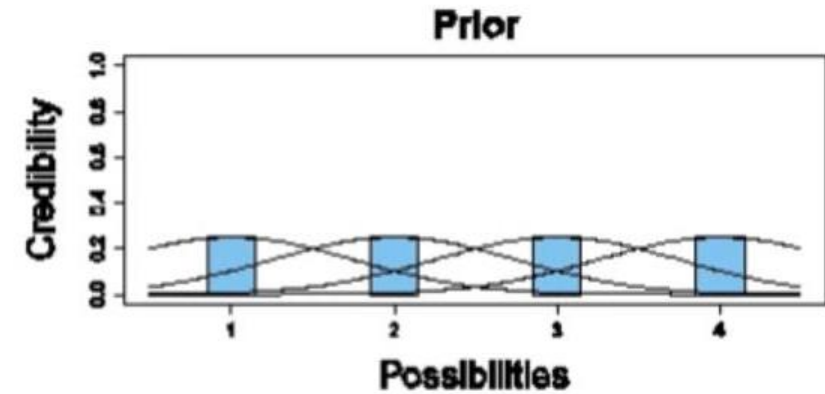
Momento 2: Distribuciones posteriores de los parámetros

Recolección de datos



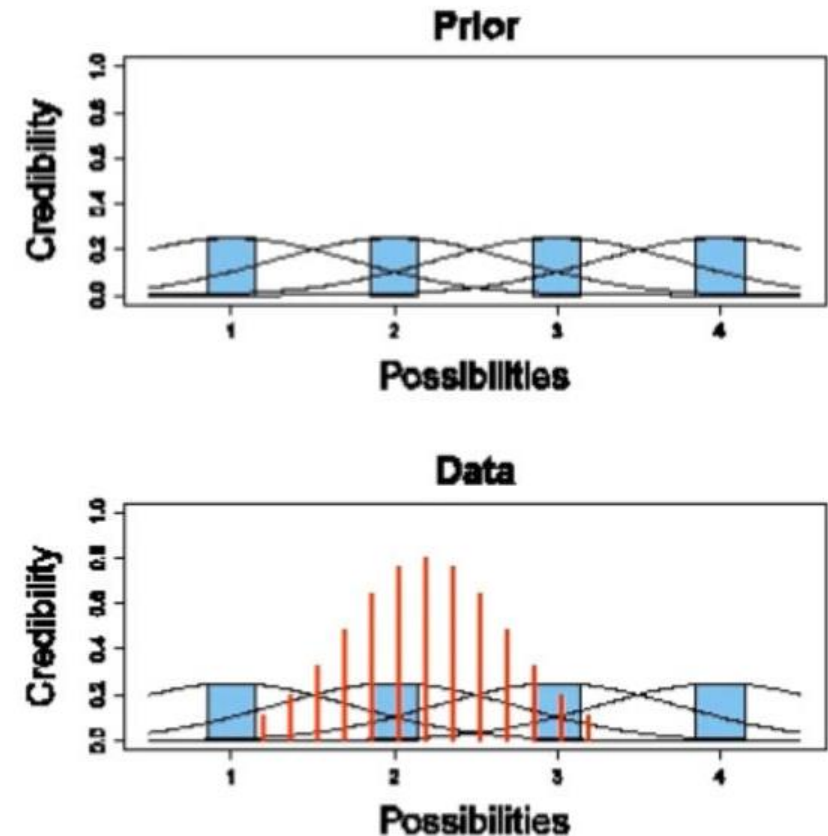
Análisis bayesiano de datos

- El papel de los datos es reubicar credibilidad entre posibilidades (i.e., aprender).
- Las posibilidades son valores de parámetros en un modelo probabilístico (μ, σ)
- Reubicamos credibilidad a valores de parámetros consistentes con los datos



Análisis bayesiano de datos

- El papel de los datos es reubicar credibilidad entre posibilidades (i.e., aprender).
- Las posibilidades son valores de parámetros en un modelo probabilístico (μ, σ)
- Reubicamos credibilidad a valores de parámetros consistentes con los datos



Análisis bayesiano de datos

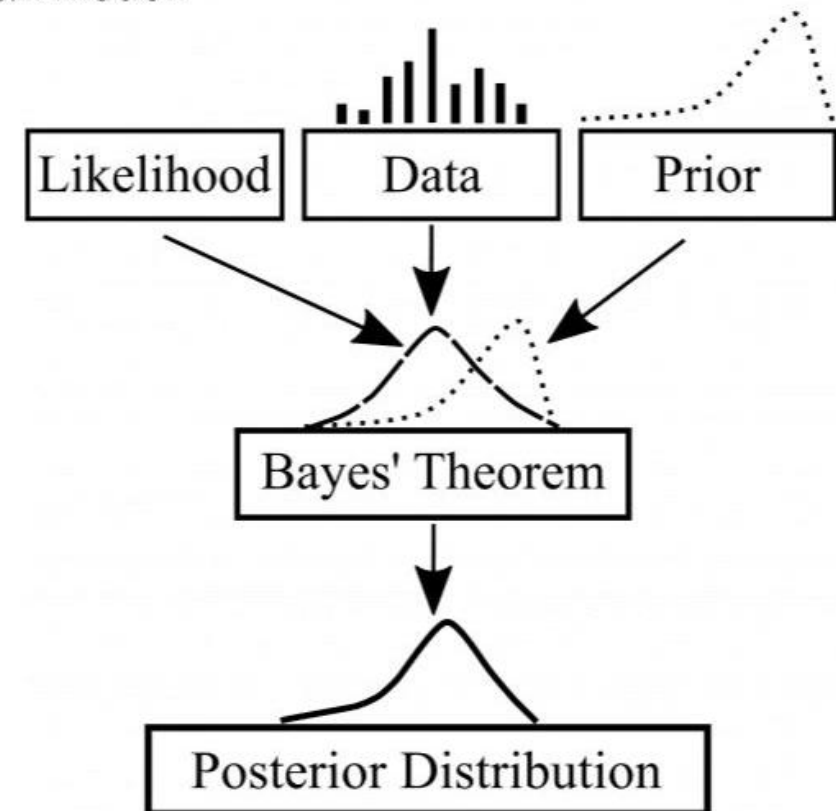
- El papel de los datos es reubicar credibilidad entre posibilidades (i.e., aprender).
- Las posibilidades son valores de parámetros en un modelo probabilístico (μ, σ)
- Reubicamos credibilidad a valores de parámetros consistentes con los datos

Formula:

$$f(\text{model} \mid \text{data}) = \frac{f(\text{data} \mid \text{model}) \times f(\text{model})}{f(\text{data})}$$

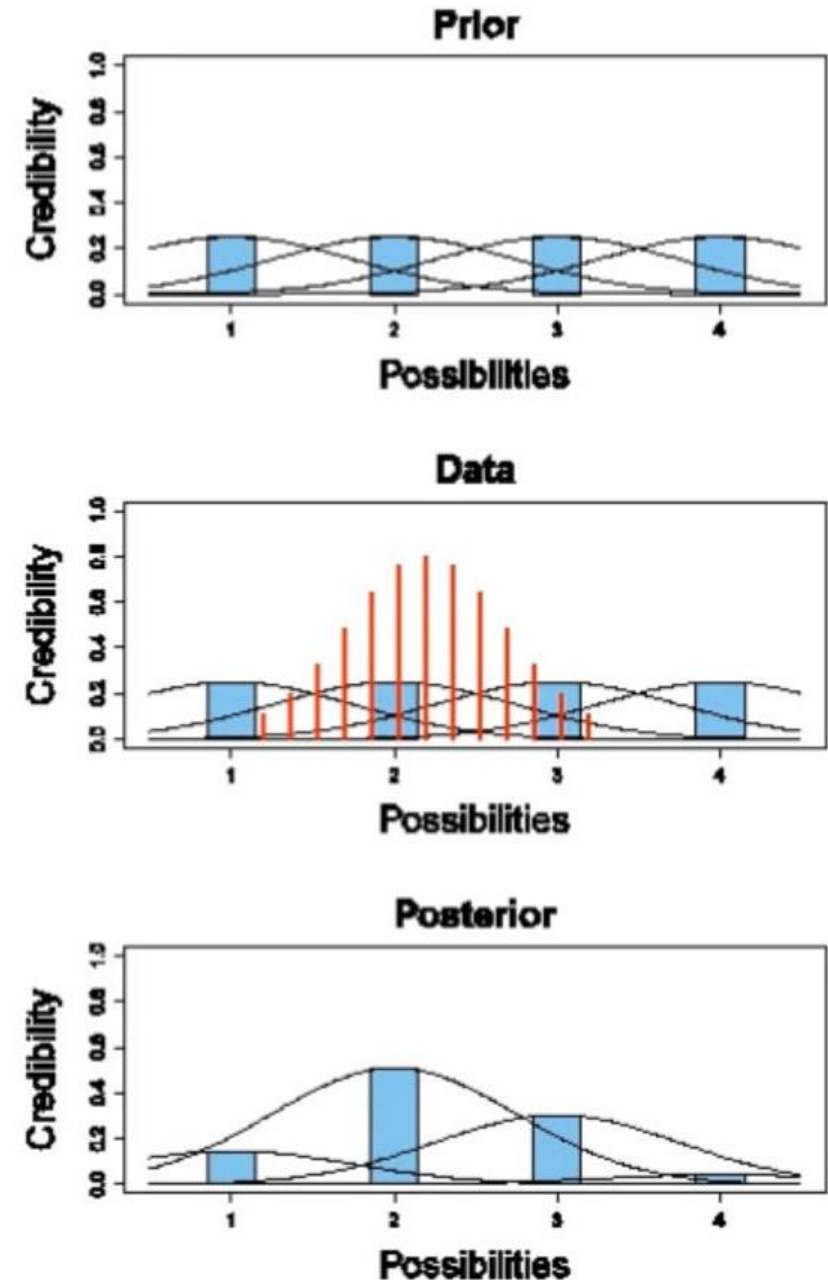
Bayesian Model Likelihood distribution Prior distribution Data

Bayesian Model:

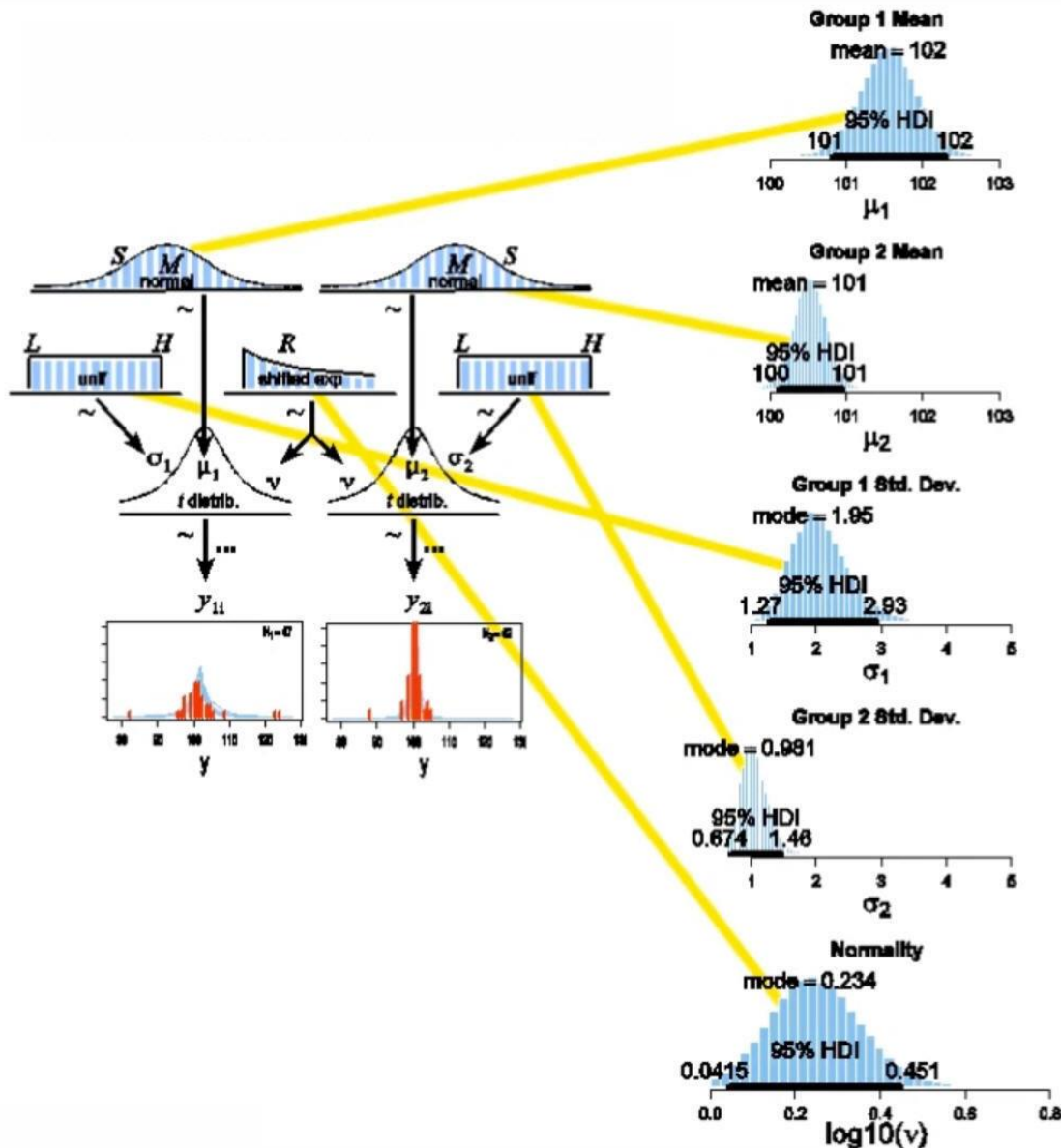


Análisis bayesiano de datos

- El papel de los datos es reubicar credibilidad entre posibilidades (i.e., aprender).
- Las posibilidades son valores de parámetros en un modelo probabilístico (μ, σ)
- Reubicamos credibilidad a valores de parámetros consistentes con los datos



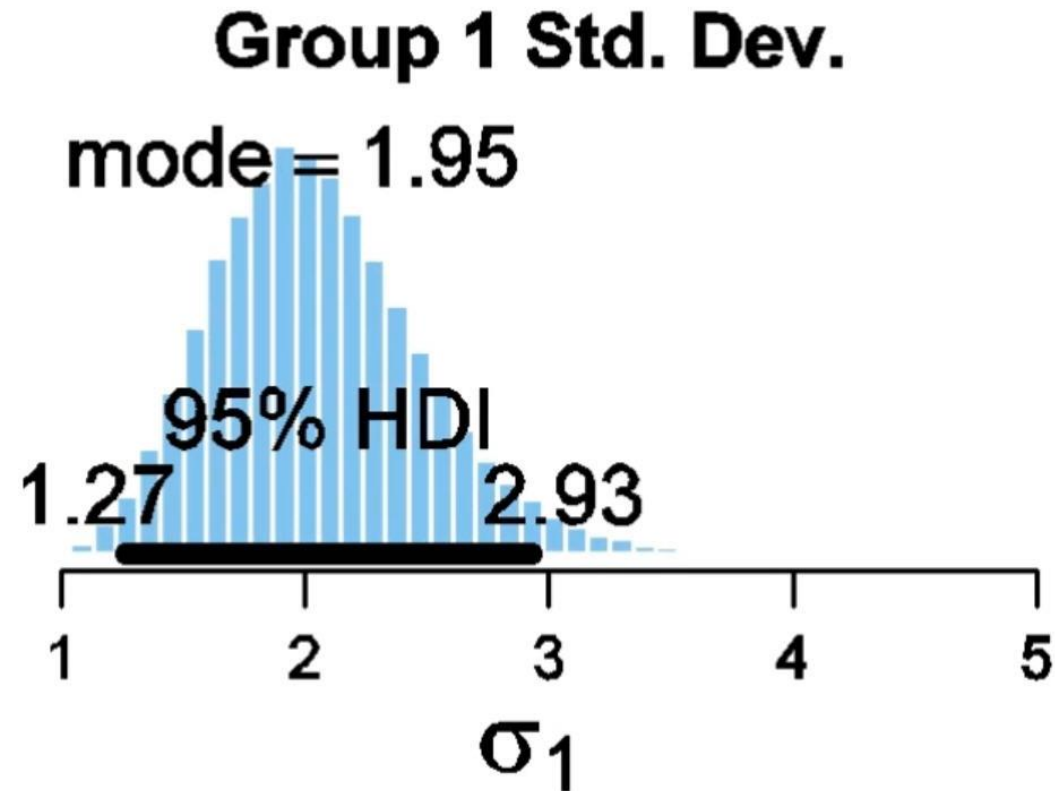
Calcular la posterior de los parámetros (MCMC)



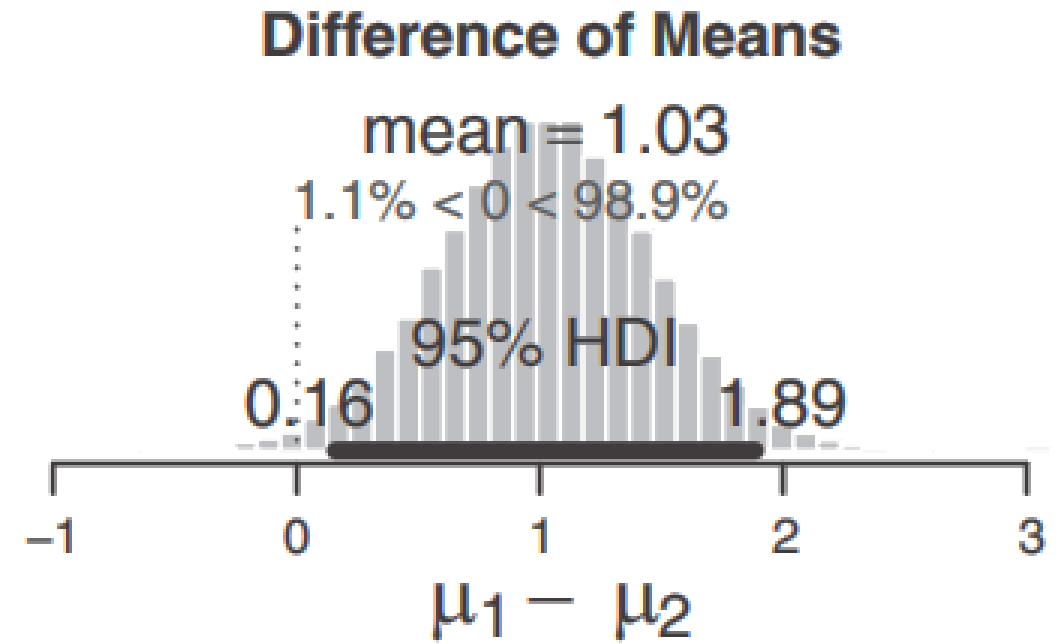
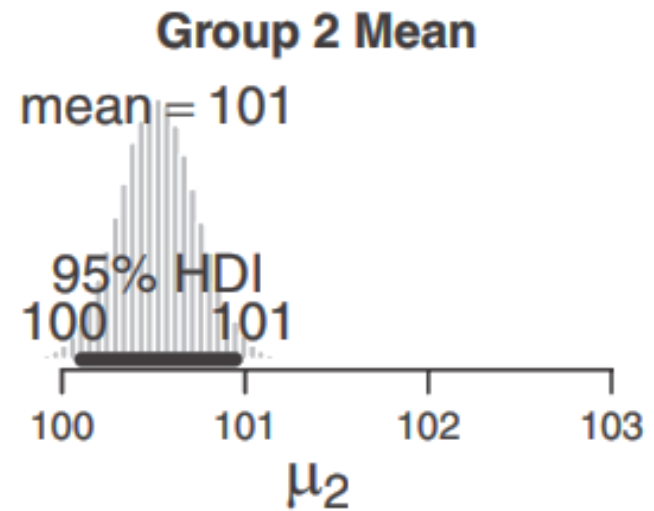
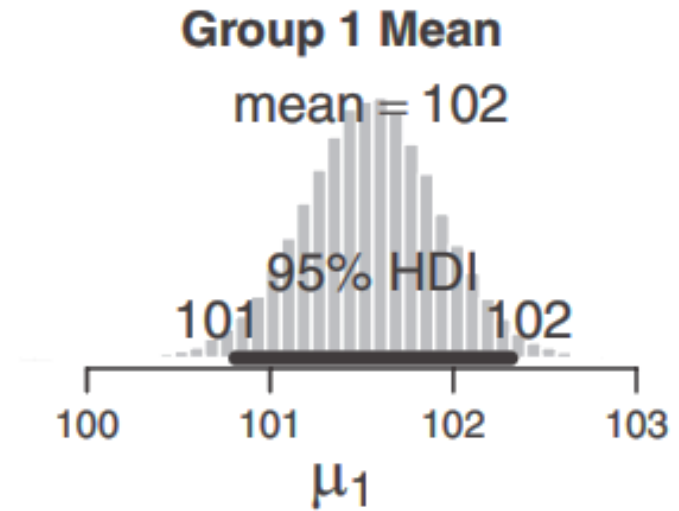
- Estos son histogramas de los valores de parámetros de la distribución posterior: un número grande de combinaciones de μ_1 , μ_2 , σ_1 , σ_2 , ν que son creíbles en conjunto dados los datos
- No son distribuciones de los datos, y no son distribuciones muestrales dada una hipótesis nula

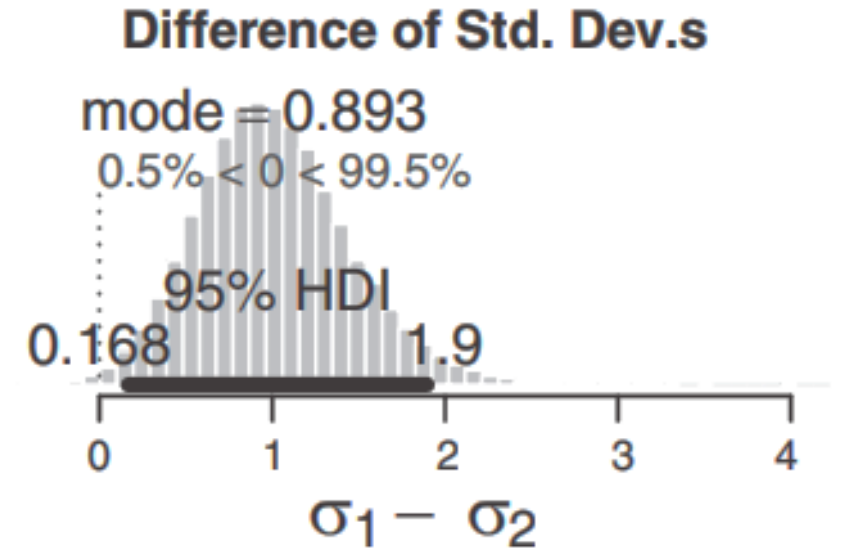
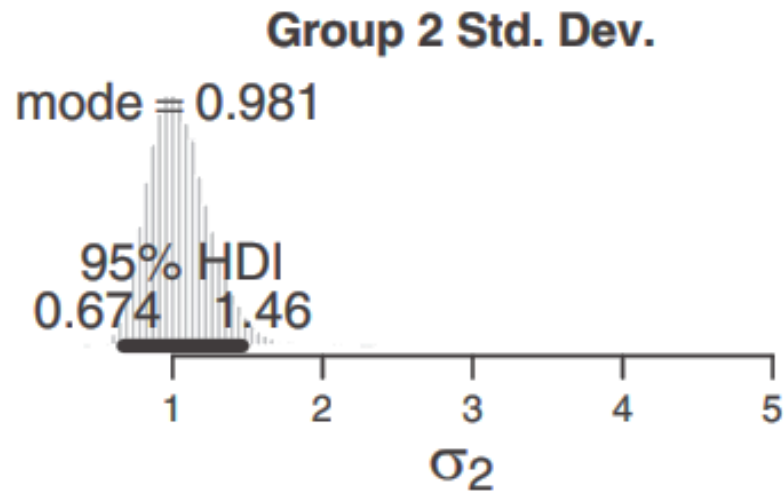
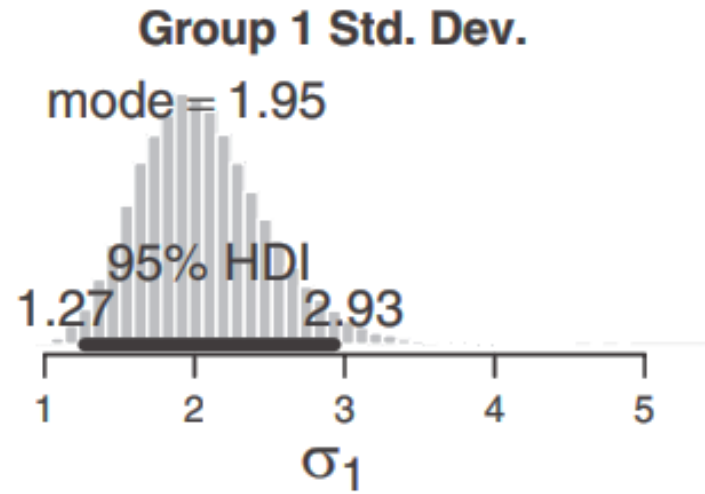
Intervalos de mayor densidad

95% HDI: Highest Density Interval



Distribuciones posteriores de los parámetros





Distribuciones posteriores de los parámetros

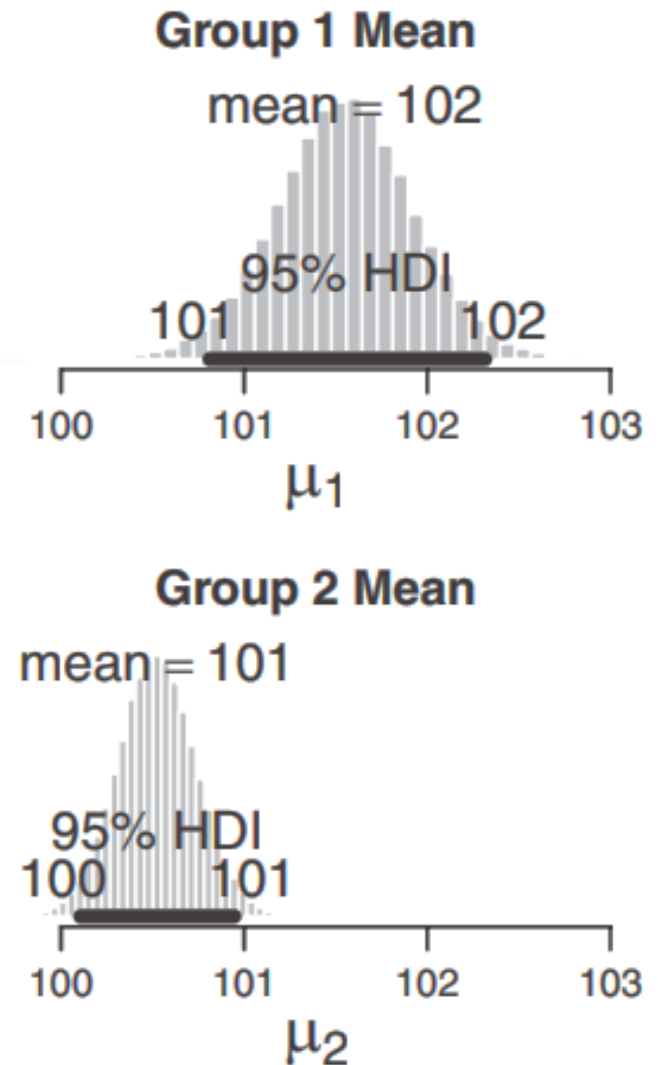
¿El t-test clásico/frecuentista hace lo mismo cierto?



Golem sin modelo generativo

¿El t-test clásico/frecuentista hace lo mismo cierto?

- Para poder hacer cualquiera de las siguientes afirmaciones:
 - Es muy probable (95%) que las medias del grupo 1 y 2 sean diferentes
 - La magnitud de la diferencia es de 1.03 con un rango creíble de .4-1.8
- Se necesita la distribución de los valores probables de μ_1 & μ_2



¿Qué tenemos entonces?

```
> print(t_test_result)

      Two Sample t-test

data:  value by group
t = 1.5587, df = 87, p-value = 0.1227
alternative hypothesis: true difference in means between group Drug and group Placebo is not equal to 0
95 percent confidence interval:
 -0.428653  3.544155
sample estimates:
 mean in group Drug mean in group Placebo
      101.9149      100.3571

> |
```

El t-test frecuentista no nos dice nada de la probabilidad de que las medias sean exactamente iguales o diferentes.

Lo que sí nos dice es que es probable que la diferencia observada sea producto de la suerte, dado que son exactamente iguales.

El resultado no sólo es contrario al que obtuvimos con la aproximación bayesiana sino que nos da poca evidencia para hacer la inferencia que queríamos hacer.

Sin la distribución del parámetro de interés (medida explícita), no hay inferencias ricas y razonables.

¿Prueba de significancia de hipótesis nula?

Si usamos una prueba por remuestreo

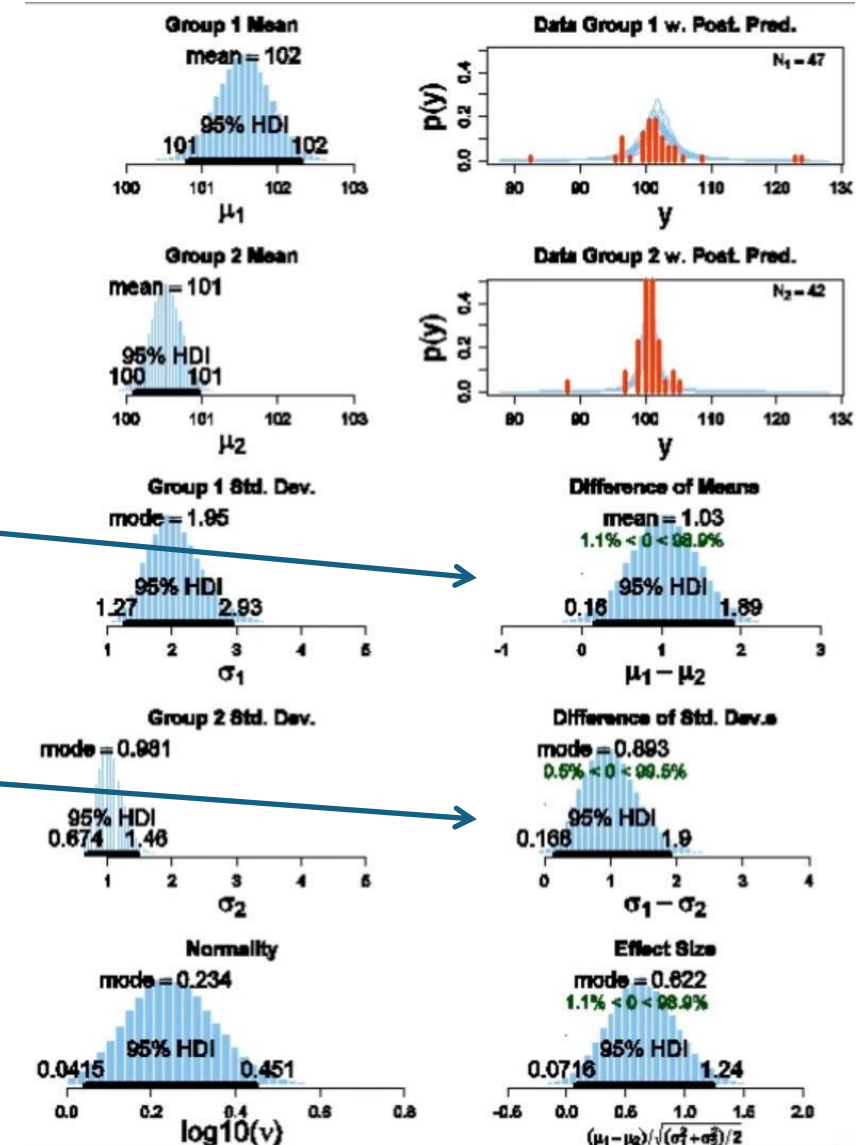
- La prueba de diferencia de medias:

$$p = 0.116(> .05)$$

- La prueba de diferencia de desviaciones estándar:

$$p = 0.072(> .05)$$

Y, debe hacer correcciones por múltiples pruebas



Pero el t-test es un modelo particular

- Todos los modelos estadísticos tienen parámetros
- No podemos saber “la verdad”; por eso describimos la incertidumbre sobre el valor de los parámetros con distribuciones
- Es sobre dichas distribuciones que busco hacer inferencias sobre la credibilidad de sus valores
- La estimación bayesiana rinde inferencias más ricas

Conclusiones

- Preguntas científicas implican algún tipo de generalización
- Tal generalización precisa medidas explícitas y congruentes sobre la credibilidad de los parámetros de un modelo
- Los parámetros tienen dos momentos en la inferencia bayesiana y lo que busco es, con base en datos, reubicar la credibilidad de los valores de dichos parámetros



Nos interesan modelos estadísticos justificados por modelos generativos

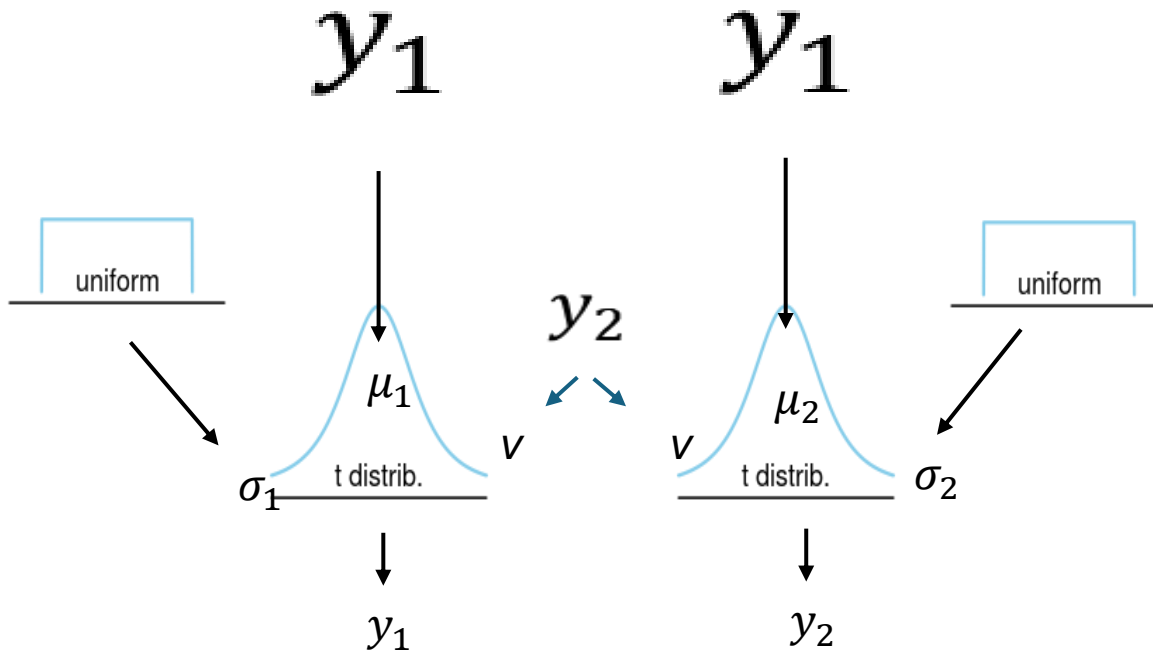
Conclusiones

- La hipótesis nula y sus tests no me brindan medidas explícitas ni congruentes para problemas científicos
- Quizá sí de control de calidad en industrias
- Los modelos nulos no existen en los estudios observacionales
- La equiprobabilidad no es una teoría sino la carencia de teoría

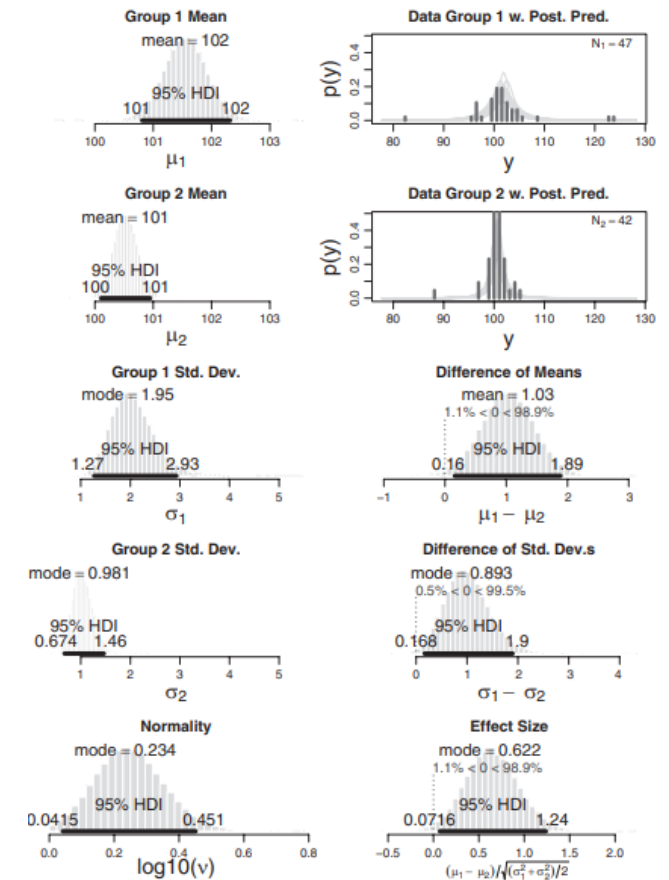
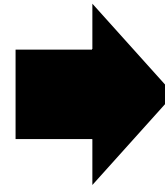


Hacia adelante: ¿Bajo qué reglas reubico credibilidad?

$$P(\Theta | D)$$

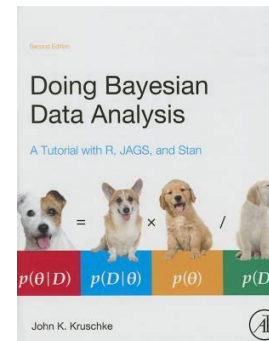


¿Reubicación de credibilidad?



Siguiente clase

- Capítulo 2 Small worlds and large worlds. Statistical Rethinking
- Capítulo 5. Bayes rule.



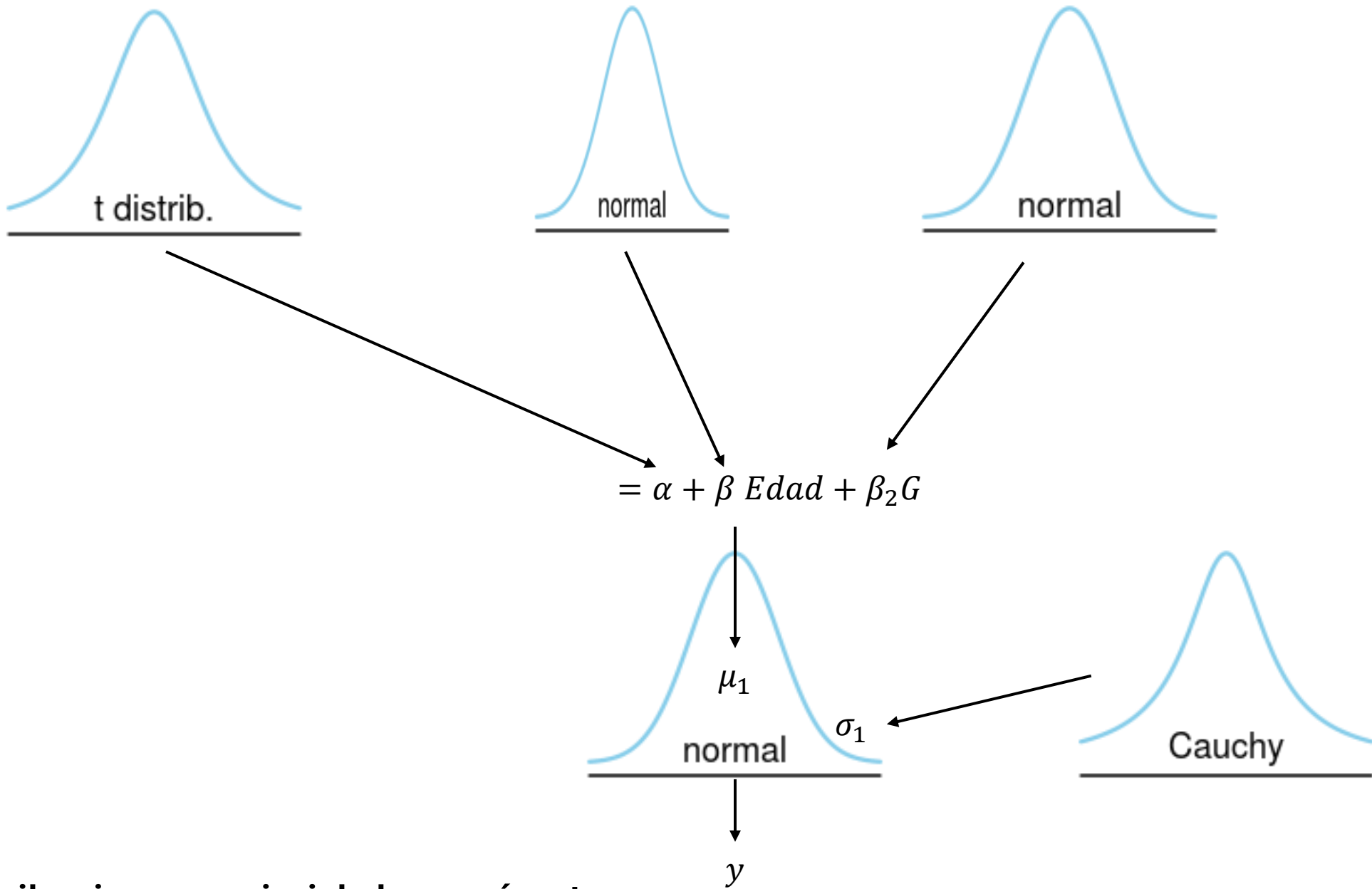
Por ejemplo

¿Qué tan distinto es un grupo respecto a otro?

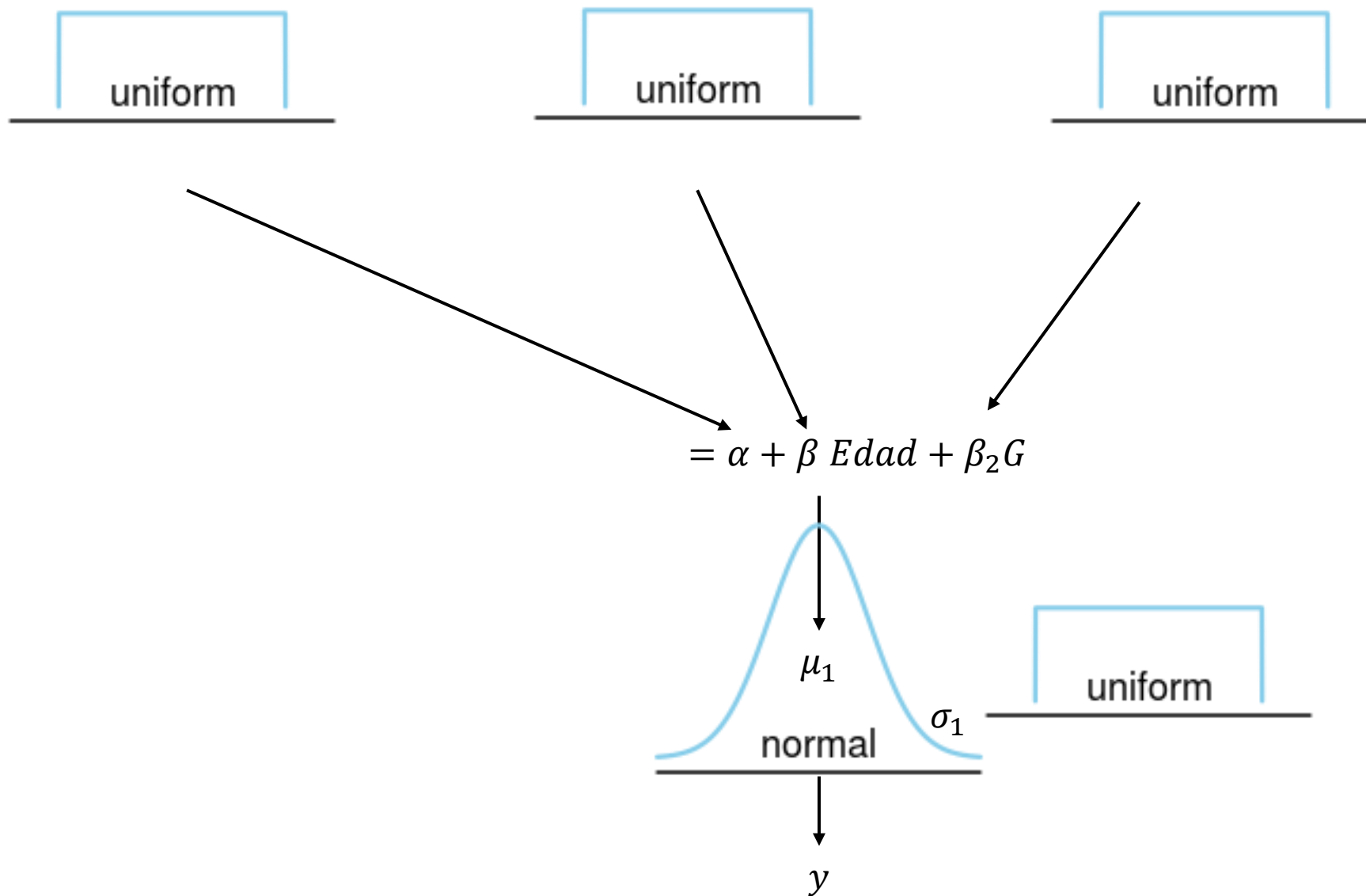
H: El salario promedio del grupo 1 es mayor al del grupo 2
condicionando por edad

Si tomo una muestra de ambos grupos:

¿Basta con comparar aritméticamente las distribuciones de los
datos observados para poder **generalizar** mis hallazgos respecto a
H?



Distribuciones a priori de los parámetros



Distribuciones a priori de los parámetros: El modelo “determinístico” es el mismo pero el modelo generativo no

Reasignación de credibilidad vía distribuciones de parámetros

- Información de que el trabajo de campo fue irregular y se tradujo en error sistemático

