Introducción a los métodos estadísticos bayesianos en Ecología

Pablo Inchausti

Programa del curso:

1. Introducción general
2. Elementos básicos del análisis bayesiano
3. Análisis bayesiano I
4. Análisis bayesiano II
5. Modelos bayesianos jerárquicos

1. Introducción general: Teórico 01

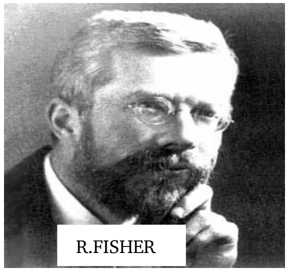
a. Dualidad en la Estadística actual.
b. Historia de la Estadística en 6 diapos.
c. El teorema de Bayes.
d. La curiosa historia de la estadística que aprendieron en otros cursos.
e. La regla de Bayes en estadística.
f. ¿Por qué no "somos todos bayesianos"?

a. Dualidad en la Estadística actual:

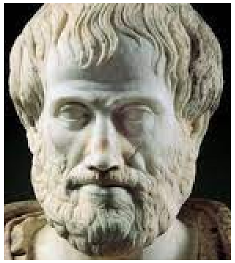
Las bases conceptuales de la Estadística actual se formularon en 1920s y 1930s como nuevas formulaciones o como rechazo a métodos previos.

En la Estadística, coexisten dos marcos teóricos con métodos diferentes de análisis desde sus orígenes. Frecuentista y Bayesiana.

Las diferencias fundamentales tienen que ver con la visión de la probabilidad y con el uso de la información previa en el análisis de datos.



Tuvo un rol fundamental en el rechazo de las métodos bayesianos y en el desarrollo de la estadística frecuentista.



definió las bases conceptuales de la lógica que permite deducir todas las conclusiones válidas y ciertas a partir de los axiomas.

No era necesaria la validación, opinión o percepción externa para obtener conocimiento verdadero.

Revolución científica: la observación y la inducción como bases fundamentales de la búsqueda de conocimiento empírico.

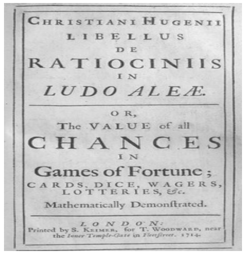
Las causas reales de los fenómenos observables: de deducidas y ciertas → inferidas e inciertas.

La probabilidad emergió como la forma de actuar bajo incertidumbre con en un juego de azar.

Los juegos de azar son de vieja data, y han tenido maravillosas explicaciones...

b. Historia de la Estadística en 6 diapos:

La teoría de probabilidad creada por B.Pascal y P.Fermat en 1654 cuantifica la incertidumbre de los eventos o resultados plausibles.



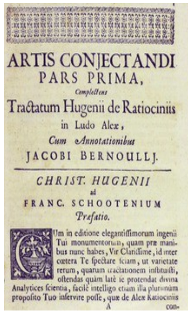
CHRISTIANI HUGENII
LIBELLUS
DE
RATIOCINIIS
IN
LUDO ALEÆ.
OR,
The VALUE of all
CHANCES
IN
Games of Fortune;
CARDS, DICE, WAGERS,
LOTTERIES, &c.
Mathematically Demonstrated.

LONDON:

Se publica el 1er libro sobre teoría de probabilidad en 1657, aún solo aplicada a los juegos de azar.

Christiaan Huygens

J. Bernoulli (1713) define la probabilidad en [0,1], deduce la ley de grandes números, introduce las nociones de probabilidad frecuentista y subjetiva y realiza aplicaciones en problemas donde hay incertidumbre de decisión.



ARTIS CONJECTANDI
PARS PRIMA,
Tractatum Hugenii de Ratiociniis
in Ludo Aleæ,
Cum Annotationibus
JACOBI BERNOULLI.
CHRIST. HUGENII
FRANC. SCHOOTENIUM
Præfatio.

Las dos visiones diferentes de la probabilidad son:

En la visión aleatoria,

Pr → la frecuencia relativa de ocurrencia de un evento en un número infinito de realizaciones.

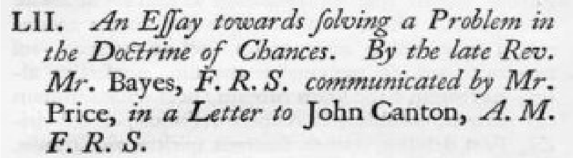
Pr(cara)=1/2 → se espera obtener "cara" en la mitad de muchas realizaciones potenciales del experimento.

En la visión epistémica,

Pr → medida del estado de conocimiento personal que refleja la plausibilidad de ocurrencia de un evento (degree of belief).

Pr(cara)=1/2 → la creencia en la plausibilidad de un resultado específico basada en suposiciones personales y/o observaciones previas.

P. Laplace (1812) formaliza las nociones de probabilidad, teorema central del límite, probabilidad condicional. **C. Gauss** "redescrubre" la dist. Normal y los mínimos cuadrados.



LII. *An Essay towards solving a Problem in the Doctrine of Chances. By the late Rev. Mr.* Bayes, *F. R. S. communicated by Mr.* Price, *in a Letter to* John Canton, *A. M. F. R. S.*

Thomas Bayes (1763) publica su hoy famoso teorema o regla.

Bayes dedujo $Pr(A|B) * Pr(B) = Pr(B|A) * Pr(A)$ a partir de las probabilidades de eventos condicionales.

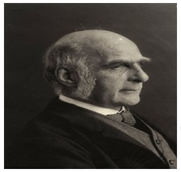
↳ (a veces llamada "ley de probabilidad inversa")

$$Pr(hipótesis|efecto) * Pr(efecto) = Pr(efecto|hipótesis) * Pr(hipótesis)$$

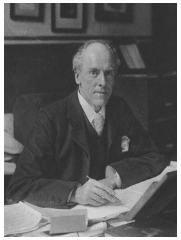
Podemos calcular la probabilidad de que una hipótesis sea cierta dado un efecto observado.

Pero fue **Laplace (1812)** quien independientemente dedujo con rigor la forma actual de regla de Bayes.

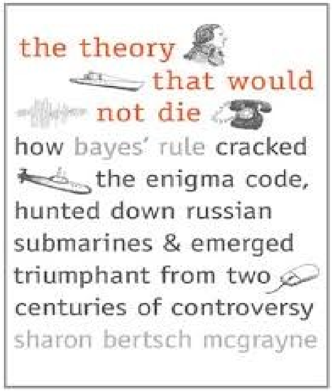
G. Achenwall acuñó en 1749 “statistisk” para referirse a la información de interés para los estados: # de nacimientos y muertes, de prisioneros acusados y condenados, de suicidios, composición etaria, tamaño de propiedades, peso de las cosechas, etc.

A. Quetelet y F. Galton emblematizan el interés en la recolección y análisis de información empírica en el siglo XIX usando los métodos desarrollados por Gauss y Laplace.



Inspirado por Galton, K. Pearson formalizó la relación entre la teoría de probabilidades y la descripción de la información empírica, creándose la estadística moderna.



Hasta la creación de la estadística moderna hacia 1920s, la mayoría de los (pocos) análisis de datos realizados tenían en una perspectiva estrictamente Bayesiana. ¿qué es "clásico"?

Como veremos: $Pr\left(posterior\right) \approx Likelihood * Pr\left(previa\right)$

Cómo definir las dist. previa fue y continúa siendo un aspecto polémico en estadística bayesiana.

Sólo algunas distr previas (conjugadas) admiten una solución analítica; para las otras, simulación.

Consideró la distr previas un aspecto arbitrario y subjetivo (lo es) a eliminar para desarrollar métodos de análisis de datos objetivos y reproducibles.

Ordenando los principales eventos en una cronología:

Century	XVII	XVIII	XIX	XX		
Theory of probability	1654 B. Pascal & P. Fermat exchange 1657 C. Huygens' book	1713 J. Bernoulli's book	1866 J. Venn book	Books by 1921 J. Keynes 1925 F. Ramsey 1933 B. de Finetti 1933 A. Kolmogorov		
Frequentist Statistics				1922,1925,1935 R. Fisher books and papers	1932, 1934 J. Neyman & E. Pearson papers	
Bayesian Statistics		1763 T. Bayes paper	1812 P. Laplace book	1939 H. Jeffries book	1953 Metropolis et al paper	1990s Rediscovery of MCMC

Este redescubrimiento del algoritmo MCMC permitió "independizar" a los métodos Bayesianos de las soluciones analitícas y es responsable de su auge actual.

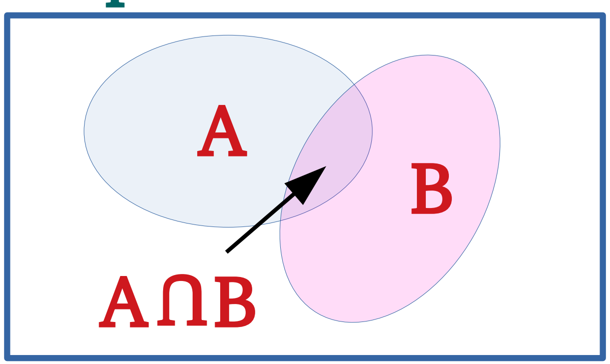
c. El teorema de Bayes.

Solo para sufrir un poco, recordemos las reglas básicas de cálculo de probabilidades ya aprendidas.

Si A y B son eventos independientes, $Pr(A \cap B) = Pr(A) * Pr(B)$

Pero si NO fueran eventos independientes, la ocurrencia de A cambia la Pr ocurrencia de B.

Espacio muestral



Como estos eventos están relacionados, la ocurrencia de A indica que su complemento ("no A") no ocurrió. Ello restringe el espacio muestral de B.

Por ende, la Pr(B|A ya ocurrió) es:

$$Pr(B|A) = \frac{Pr(A \cap B)}{Pr(A)}$$

El teorema de Bayes se obtiene a partir de las Pr. de eventos condicionales y simplemente dice:

$$Pr(A|B)*Pr(B)=Pr(B|A)*Pr(A)$$

El teorema de Bayes es una regla para calcular la "probabilidad inversa": conociendo Pr(A) y Pr(B), dado Pr(A|B), podemos ahora calcular Pr(B|A).

Una aplicación: detección de cáncer prostático.

Un paciente recibe un test PSA+:

¿cuál es la probabilidad de que esté enfermo?

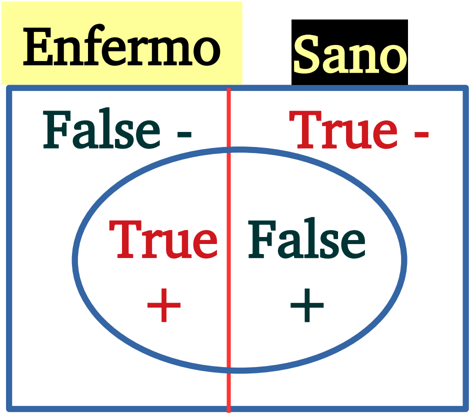
Sean A: estado del paciente, B: resultado del test.

$$Pr(Enfermo|+)*Pr(+)=Pr(+|Enfermo)*Pr(Enfermo)$$

Lo que quiere el paciente

Lo que dice el test

Incidencia enfermedad



Sensibilidad = True+|enfermo

Especificidad = True-|sano

Propiedades del Test PSA:

sensibilidad~ 0.35 y especificidad~0.15.

→ 65 de cada 100 tests + son falsos.

→ 85 de cada 100 tests - son falsos.

A partir de la regla de Bayes se obtiene:

$$Pr\left(Enfermo|+\right)=\frac{Pr\left(+|Enfermo\right)*Pr\left(Enfermo\right)}{Pr\left(+\right)}$$

Lo que quiere el paciente

Lo que dice el test (sensibilidad)

Incidencia enfermedad.

Pr(+): incluye a los True+ y False +

Antes de examinar a ese paciente, ¿qué sabemos de la enfermedad en la población?

Pr(+): todas las formas en que se puede obtener como resultado un test +: True+ y False+.

$$Pr(+) = Pr(+|Enfermo) * Pr(Enfermo) + Pr(+|Sano) * Pr(Sano)$$

Substituyendo:

$$Pr(Enfermo|+) = \frac{Pr(+|Enfermo) * Pr(Enfermo)}{\underbrace{Pr(+|Enfermo) * Pr(Enfermo)}_{\text{True +}} + \underbrace{Pr(+|Sano) * Pr(Sano)}_{\text{False +}}}$$

Para calcular Pr(Enfermo), hay que conocer la incidencia de la enfermedad en la población a la que pertenece el paciente.

¿Qué sabemos del cancer prostático en Uruguay?

*** solo para hombres * solo mayores de 40 años.**

$$Incidencia\ en\ Uruguay = \frac{1456\ casos\ por\ año}{3.300.000 * 0.5 * 0.4} = 0.22\,\% \sim \frac{1}{453}$$

$$Pr(Enfermo|+) = \frac{Pr(+|Enfermo) * Pr(Enfermo)}{Pr(+|Enfermo) * Pr(Enfermo) + Pr(+|Sano) * Pr(Sano)}$$

$$Pr(Enfermo|+) = \frac{0.35 * \left(\frac{1}{453}\right)}{0.35 * \left(\frac{1}{453}\right) + (1 - 0.15) * \left(\frac{452}{453}\right)} = 0.09\% \sim \frac{1}{1098}$$

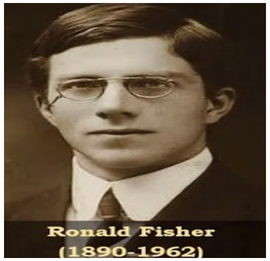
(hay maneras más simples de hacer los mismos cálculos)

Lo que importa aquí es entender que la regla de Bayes permite calcular la "probabilidad inversa":

i.e. dado el "efecto observado" (resultado + del test), calcular la "probabilidad de la causa" (enfermo).

Para las ♀ quienes carecen de próstata: la sensibilidad y especificidad del mamograma son 0.87 y 0.85.

d. La curiosa historia de la estadística que aprendieron en otros cursos.



publicó entre 1922 y 1935 una serie de artículos y libros que cambiaron la teoría y sobre todo la práctica de la estadística.

En ellos:

* Distribuciones muestrales (junto con “Student”) permitiendo inferencia con muestras “pequeñas”.

* Máximo de verosimilitud: método genérico de estimación de parámetros e inferencia.

* Test de hipótesis estadísticas y el p-valor.

* Escribió dos libros fundamentales: Statistical methods for research workers (1925) y Design of experiments (1935).

Fisher (1925) desarrolló el test de significancia estadística para proveer un método objetivo para realizar inferencia inductiva sobre una sola hipótesis a partir de resultados experimentales.

Se basa en la ley de contraposición (modus tollens).

Para decidir sobre la veracidad de LA hipótesis, Fisher propuso el p-valor: Pr(resultados iguales o más extremos a los obtenidos|Hipótesis es cierta)

El p-valor NO es una prueba probabilística de la contradicción porque Pr(datos extremos|hipótesis es cierta)≠Pr(hipótesis es cierta|datos extremos)

$$\Pr(\underbrace{\text{ser católico}}_{\text{OBS}}|\underbrace{\text{es el papa}}_{\text{hipótesis}}) \neq \Pr(\underbrace{\text{ser el papa}}_{\text{OBS}}|\underbrace{\text{es católico}}_{\text{hipótesis}}).$$

El p-valor supone a la hipótesis como cierta, y así su veracidad/falsedad no es evaluada directamente.

Fisher defendió encarnizadamente la visión frecuentista de la probabilidad. Para él, los datos (y los estadísticos) eran una realización posible de extraer una muestra aleatoria de una población estadística de tamaño infinito.

p-valor=0.02: sólo en 2 de 100 muestras aleatorias potenciales se obtendrían resultados iguales o más extremos si la hipótesis fuese cierta.

Interpretación del p-valor: medida continua de compatibilidad de la información muestral con la hipótesis que se está evaluando.

¿Cómo transformar p-valor=0.02 en un criterio binario de rechazo/no rechazo la hipótesis?:

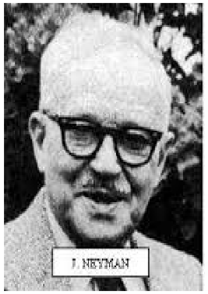
Fisher (1925 p.47) sugirió que era "conveniente en este momento usar P = 0.05 como un límite para juzgar cuando una desviación puede ser considerada significativa o no".

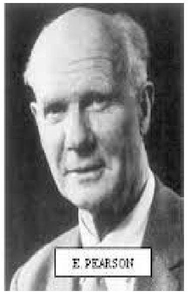
¿Por qué 0.05?

COMMERCIAL DISPUTE

Porque K.Pearson le negó el uso de sus extensas tablas publicadas y Fisher recalculó solo algunas...

Este innovador procedimiento de inferencia para hipótesis específicas (ω=0.14) fue criticado por carecer de otras hipótesis y por su falta de rigor matemático para decidir el estadístico del test a emplear para calcular el p-valor.





empezaron (1928) tratando de mejorar el método de Fisher y terminaron (1933 y 1934) creando un enfoque alternativo.

Su idea era definir H_0 y H_1 específicas ($\omega=0.14$ vs $\omega=0.42$) y decidir con test de significancia si datos apoyan a H_0 ó a $H_1 \leftarrow$ teoría de la decisión.

La existencia de 2 hipótesis específicas implicó:

- $\alpha=\Pr(\text{Rechazar } H_0 | H_0 \text{ cierta})$
- $\beta=\Pr(\text{Rechazar } H_1 | H_1 \text{ cierta})$
- Magnitudes de efecto
- Potencia del test.

Lema Neyman-Pearson: el LRT de las dos hipótesis tiene máxima potencia.

N-P requerían predefinir el tamaño muestral para obtener la potencia deseada SEGÚN EL CASO.

Si bien tenía 2 hipótesis específicas, solo la "hipótesis principal" se usaba para en la decisión!

El error de decisión a minimizar (**α** ó **β**) variaba SEGÚN EL CASO en función de las implicaciones.

Cuando se rechazaba una hipótesis para un α pre-fijado, la otra se "aceptaba por implicación".

"Mecanización del proceso de decisión": rechazar hipótesis si p-valor < **α**.

Interpretación del resultado: la inferencia sólo posible como "comportamiento inductivo" a largo plazo – serie de decisiones en contexto idéntico.

El procedimiento NP de decisiones repetidas es más apropiado en un contexto de control de calidad, no de experimentos científicos donde jamás hay condiciones idénticas de repetición.

La formulación de 2 hipótesis específicas y las diferencias de interpretación desataron una polémica que solo terminó cuando Fisher murió!

VOLUME 3 — *March 1961* — NUMBER 4

SILVER JUBILEE OF MY DISPUTE WITH FISHER[1]

JERZY NEYMAN

<u>Fisher to Neyman</u>: "a threat to the intellectual freedom of the West(ern Hemisphere)"

<u>Neyman to Fisher</u>: "worse than useless"

Fisher y Neyman solo estaban de acuerdo en que a) los métodos bayesianos eran inaceptables, b) el enfoque del otro era totalmente equivocado.

Desde la posguerra, los libros de texto de estadística hicieron una "síntesis impersonal" de los enfoques de Fisher y Neyman-Pearson.

Este "algoritmo híbrido" pretendidamente objetivo basado en muestras potenciales de una población infinita o en repeticiones idénticas del mismo experimento que uds y yo aprendimos consiste en:

1) Definir dos hipótesis mutuamente excluyentes. H_0:"no efecto" vs H_1:"algún efecto".

2) Suponiendo H_0 cierta, calcular estadístico (T).

3) Comparar T con un cuantil C de FuncDensProb (suponiendo H_0 cierta) para $\alpha=0.05$ y gl dados.

4) Si el valor de T fuese más extremo que C (o el p-value $<\alpha=0.05$) → rechazar H_0.

Mecánico, objetivo, impersonal e irreflexivo: perfecto para implementarlo en una máquina.

Hay ríos de tinta vertidos y miles de páginas escritas acerca de los errores, e inconsistencias de la inferencia frecuentista y las diferencias entre Fisher y Neyman-Pearson desde 1930.

EL MUNDO AL REVÉS

<u>Lo que obtenemos:</u>

Pr (“datos”$|H_0$ cierta)

Estadística frecuentista

<u>Lo que quisiéramos:</u>

Pr(R H_0?$|$evidencia)

Estadística bayesiana

Pero la simplificación del manual “how to...” ha sido francamente irresistible desde 1950...

e. La regla de Bayes en estadística.

Hace 11 diapos se estimó Pr(Enfermo|Test+) a partir de un resultado (Test +), la sensibilidad (Pr (Test+|Enfermo)) y especificidad (Pr (Test-|Sano)) y la incidencia de la enfermedad (Pr(Enfermo)).

$$Pr(Enf|+) = \frac{Pr(+|Enf)}{Pr(+|Enf)*Pr(Enf)+Pr(+|Sano)*Pr(Sano)}*Pr(Enf)$$

Más generalmente, la probabilidad de cada observación Y_i se calcula con una FuncDensProb.

Ej. sean datos de conteos y Y ~ Poisson(μ_Y).

La inferencia (y las hipótesis) se refiere al valor del parámetro μ_Y que en el modelo Poisson permite predicir los valores esperados de Y.

En consecuencia, la ecuación se transforma en:

$$Pr(\mu_Y|datos) = \frac{Pr(datos|\mu_Y)}{Pr(datos)} * Pr(\mu_Y)$$

$Pr(\mu_Y|datos)$: prob. de obtener cada valor μ_Y DADOS LOS DATOS OBSERVADOS **POSTERIOR**

$Pr(\mu_Y)$: prob. de cada valor μ_Y ANTES DE HABER OBTENIDO LOS DATOS OBSERVADOS. **PREVIA**

$Pr(datos)$: prob. de obtener los datos considerando todos los valores posibles de μ_Y.

$Pr(datos|\mu_Y)$: prob. de obtener los datos observados dado el valor de μ_Y. **VEROSIMILITUD**

$$\boxed{Pr(\mu_Y|datos)=\frac{Pr(datos|\mu_Y)}{Pr(datos)}*Pr(\mu_Y)}$$

Dos observaciones claves:

1) La verosimilitud Pr (datos|μ_Y) se refiere al valor de μ_Y que más probablemente podría haber generado los datos observados.

El método de máxima verosimilitud es también usado en estadística frecuentista.

2) Para resolver el problema, Pr(datos) no es importante ya que solo es necesario obtener la "verosimilitud relativa".

(Ya que $\mu_Y \in \mathbb{R}$)

De lo contrario, habría que:

$$Pr(\mu_Y|datos)=\frac{Pr(datos|\mu_Y)}{\int Y*Poisson(Y,\mu)\,d\mu}*Pr(\mu_Y)$$

Dist posterior **Verosimilitud** **Dist previa**

$$\boxed{Pr(\mu_Y|datos) \approx Pr(datos|\mu_Y) * Pr(\mu_Y)}$$

Evidencia actual **Información en los datos** **Evidencia previa**

La regla de Bayes $Pr(\mu_Y|datos) \approx Pr(datos|\mu_Y) * Pr(\mu_Y)$ **puede interpretarse como un "procesador de información" que progresivamente refina el conocimiento sobre μ_Y previamente adquirido.**

En la estadística frecuentista, se estima el valor de μ_Y que más probablemente podría haber generado los datos observados.

En la estadística bayesiana, se obtiene una distr. de probabilidades de valores plausibles de μ_Y dados los datos y la evidencia previa.

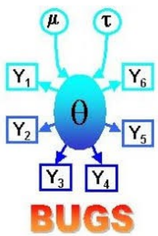
En la perspectiva frecuentista se pregunta:
"¿Cómo datos potenciales a ser obtenidos permiten discernir hipótesis (rechazar H_0)?"

En la perspectiva Bayesiana se pregunta:
"¿Cómo los datos actuales alteran la visión previa del problema (distr. previa)?"

El uso y la representación de la distr. previa fue (y aún es) durante décadas duramente debatida.

Hasta 1950s, solo habían muy pocas soluciones analíticas para la regla de Bayes para las cuales las distribución posterior cambiaban radicalmente dependiendo de la distribuciones previas empleada. ← considerado poco objetivo y potencialmente muy arbitrario.

Casi ningún modelo bayesiano tiene solución analítica posible → se requieren métodos numéricos de simulación para obtener la distr. posterior.



1989: BUGS, un programa que simplificó la aplicación del algoritmo MCMC (Markov Chain Monte Carlo; publicado en 1953).

1990s: aparecieron variantes mejoradas como WinBUGS y OpenBUGS (también para Mac y Linux)

2003: JAGS (una clara mejora sobre los BUGS).

2009: INLA
2012: Stan
(basados en otros marcos teóricos y con ciertas ventajas relativas)

(progresivamente iremos y entendiendo cómo usar el software)

Principales diferencias entre métodos frecuentistas y bayesianos y sus consecuencias:

	Estadística frecuentista	Estadística bayesiana
Visión de la probabilidad	Frecuencia relativa de un evento en un conjunto de datos potenciales.	Refleja la incertidumbre personal sobre los posibles valores de los parámetros del modelo.
Uso de la información previa	No usada explícitamente; incorporación ad-hoc a discreción del usuario.	Parte absolutamente esencial del análisis: es un "update" progresivo de información previa.
Parámetros de los modelos	Valores fijos, no observables, y no conocibles.	Valores fijos, inciertos pero estimables.
Énfasis en:	Pr(datos|hipótesis)	Pr(hipótesis|datos)
Evaluación de incertidumbre. Métodos.	Distribución muestral de réplicas potenciales. Uso de máximo de verosimilitud.	Distribución de probabilidades (posterior) obtenida por métodos basados en el teorema de Bayes usando los datos y las distr previas
Indicadores de la precisión	Intervalos de confianza	Intervalos de credibilidad

f. ¿Por qué no “somos todos bayesianos”?

Algunas ventajas de la perspectiva Bayesiana:

- Interpretación realmente intuitiva de los IC y evita las corr. Bonferroni en las comparaciones múltiples.

$$IC_{95\%}(\mu_Y) = \overline{Y} \pm t_{df,0.975} SE(\overline{Y})$$

$$IC_{95\%}(\mu_Y) = 14.7 \pm 3.2$$

Estadística frecuentista:

Si se obtuvieran k muestras aleatorias y se calcularan los IntConf, 95% de ellos contendrían el valor del parámetro.

Estadística bayesiana:

Hay una probabilidad de 0.95 que el parámetro esté en el IntCredibilidad.

(lo de las comparaciones múltiples lo explicaré después).

- Diferentes fuentes de variación se modelan de forma similar y comparables en el mismo modelo.

Toda incertidumbre es estocástica y se modela a través de una FuncDensProb para cada parámetro.

- Modelaje coherente y estimación de datos faltantes.

Los datos faltantes (que pueden ocurrir de muchas formas y con consecuencias diferentes) se tratan como un parámetro más a estimar en el modelo.

- Incorporación de errores de medición y variables latentes (i.e. no observadas) de forma coherente.

En estadística frecuentista, no hay una forma general, coherente (y comparable entre modelos) para estimar datos faltantes ni incorporar errores de medición y "estados latentes" de las variables.

- Posibilidad de incluir variación estocástica y vars. explicativas a varios niveles: modelos jerárquicos.

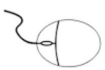
Los modelos mixtos de la estadística frecuentista, son muy limitados y poco generales en comparación...

Algunos inconvenientes de la perspectiva Bayesiana:

- Requiere mayor nivel de conocimiento en Matemática y Estadística: hay que escribir los modelos estadísticos

Muchos métodos de la estadística frecuentista han sido reducidos a “recetas simples”, automatizados en clicks de mouse y (ab)usados hasta el infinito...

Para formular los métodos bayesianos, uds deben aprender a escribir los modelos estadísticos, emplear métodos como MCMC e interpretar sus parámetros. ←

 finalmente descansará...

Ej. para una regresión Poisson, hay que escribir:

$Y \sim \text{Poisson}(\mu_Y(X))$ con $\log(\mu_Y) = \beta_0 + \beta_1 X$ cuyos parámetros β_0 y β_1 tienen distribs. previas como

$\beta_0 \sim N(\text{media}=0, \text{sd}=10)$ y $\beta_1 \sim N(\text{media}=0, \text{sd}=9.4)$

- La resolución de problemas reales es siempre numérica... y las simulaciones jamás son "exactas":

  Hay que entender un mínimo de cómo funcionan los métodos MCMC, evaluar convergencia, hacer suficientes simulaciones, chequear, etc.

- Analizar el efecto de las distribuciones previas…

  Representar la información/ignorancia previa es un aspecto delicado y discutible que se debe evaluar.

  Rehusar incorporar la información previa (¿meta-análisis?) implica no aprender nada del pasado.

- Validación de los modelos es "menos evidente"….

  Este aspecto (bien desarrollado pero poco usado en la estadística frecuentista) ha tenido gran desarrollo en los últimos 25 años...

1. <u>Introducción general</u>: Teórico 01

a. Dualidad en la Estadística actual.

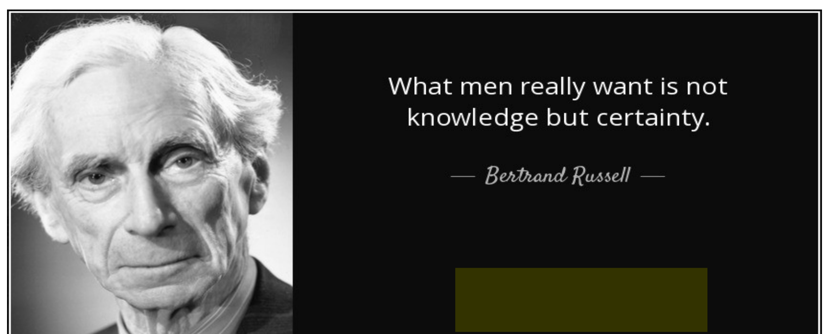
b. Historia de la Estadística en 6 diapos.

c. El teorema de Bayes.

d. La curiosa historia de la estadística que aprendieron en otros cursos.

e. La regla de Bayes en estadística.

f. ¿Por qué no “somos todos bayesianos”?



Algunas referencias:

Perezgonzalez(2014) Frontiers in Psychology 6: 223
Gigerenzer (2004) The J of Socioeconomics 33:587-606
Scheider (2014) Scientometrics 102:411–432
Fienberg (1992) Statistical Science 7: 208-225
Gigerenzer et al (1999) The empire of chance: how probability changed science and everyday life.
Cohen (1990) American Psychologist 45:1304-1312
Cohen (1994) American Psychologist 49:997-1103
Goodman (2003) Int J Epidemiology 32:699-702
Rao (1992) Statistical Science 7: 34-48
Goodman (2008) Seminars in Hematology 35:35-40
Greenland et al (2016) European J. Epidemiology 31:337-350
Colquhoun (2017) Royal Society Open Science 4: 171085

El cáncer prostático ocurre en hombres de 40+ años.

Un paciente recibe un test PSA+:
¿cuál es la probabilidad de que esté enfermo?

Para interpretar el test SE NECESITA:

- * Tasa Falso +: (1-sensibilidad) } ← test PSA
- * Tasa Falso - : (1-especificidad) }
- * Incidencia de enfermedad en la población.

Sensibilidad~ 35%

Especificidad~ 15%

Incidencia (Uruguay):

$$\frac{1456\, casos/año}{3.3*10^{6}*0.5*0.4} \sim \frac{1}{453}$$

	+	-	Total
Enfermo	0.35*1 =0.35	(1-0.35)*1 =0.65	1
Sano	(1-0.15)*452 =384.2	0.15*452 =67.8	452
Total	384.55	68.45	453

La Pr(enfermo|test+) es: $\frac{0.35}{384.55} = 0.09\,\% \sim \frac{1}{1098}$