Matias Davila
Inferencia Bayesiana

Evaluación de la Distribución Posterior en la Actividad 1A.1

En la inicial sección de nuestra actividad de inferencia bayesiana, se nos presentó el desafío de calcular una distribución posterior partiendo de una distribución muestral exponencial y una previa Gamma. Mediante la aplicación de principios bayesianos, procedí a actualizar nuestra comprensión previa del parámetro ϑ tras la observación de datos específicos. Siguiendo un enfoque metodológico riguroso, y asumiendo valores iniciales para los parámetros a y b, la distribución previa Gamma se ajustó para reflejar una nueva perspectiva posterior con parámetros actualizados a a=4 y b=10. Este proceso de actualización no solo refuerza la teoría bayesiana, sino que también ilustra la flexibilidad y adaptabilidad de la misma al integrar nueva información.

Análisis Riguroso de la Actividad 1A.2

Posteriormente, en la continuación de nuestra exploración estadística, abordé el cálculo de una distribución posterior Beta a partir de una distribución muestral geométrica y una previa Beta. Implementando la misma metodología bayesiana de actualización, los datos observados fueron esenciales para refinar la distribución previa, resultando en una distribución Beta posterior con parámetros actualizados a a=5 y b=5. Este refinamiento representa una estimación más precisa de la probabilidad de éxito ϑ, demostrando así la capacidad del análisis bayesiano para sintetizar conocimiento previo y datos empíricos en una conclusión estadística sólida y fundamentada.

```
C:\Users\matia\Documents\Matias1\Inferencia bayesiana\revaluacion>python actividad1.py
Parámetros de la distribución Gamma posterior: a=4, b=10
Parámetros de la distribución Beta posterior: a=5, b=5
```

1B

Parámetros a estimar	Distribución previa	Distribución posterior	Actualización hiperparámetros
y_i ~ Exp(θ)	Gamma(a, b)	Gamma(a + n, b + Σy)	a += n, b += Σy
y_i ~ Geom(θ)	Beta(a, b)	Beta(a + Σy, b + n)	a += Σy, b += n
y_i ~ Exp(θ)	Gamma(a, b)	Gamma(a + n, b + Σy)	a += n, b += Σy
y_i ~ Geom(θ)	Beta(a, b)	Beta(a + Σy, b + n)	a += Σy, b += n

La Actividad 1B requirió la aplicación práctica de la teoría bayesiana para construir una tabla que relacionara distribuciones muestrales con sus correspondientes previas y posteriores. Este ejercicio de síntesis no solo solidificó mi comprensión de cómo actualizar las creencias previas con datos observados, sino que también me permitió visualizar la transformación de los hiperparámetros de manera clara y ordenada. Mediante la correcta selección de distribuciones y la precisión en la actualización de parámetros, la tabla resultante refleja la esencia del razonamiento bayesiano y proporciona una herramienta útil para futuras inferencias estadísticas.

Matias Davila
Inferencia Bayesiana

Análisis de la Actividad 2a: Cálculo de la Distribución Posterior

En la Actividad 2a, abordé la tarea de actualizar la distribución previa de la tasa de llamadas en un call center, dados los datos recopilados en intervalos de dos minutos. Utilizando una distribución previa Gamma con parámetros α=2 y β=1, y datos que seguían una distribución de Poisson, calculé la distribución posterior Gamma. La suma de las llamadas proporcionó el nuevo parámetro α y el número total de datos observados ajustó el parámetro β. Como resultado, obtuve una distribución posterior Gamma con parámetros α=105 y β=51, que reflejan la actualización de nuestras creencias sobre la tasa de llamadas basada en la evidencia reciente.

Análisis de la Actividad 2b: Estimadores Bayesianos de Riesgo

En la Actividad 2b, procedí a calcular estimadores bayesianos del parámetro ϑ para diferentes funciones de riesgo: cuadrática, lineal y discreta. Utilizando la distribución posterior Gamma obtenida en la Actividad 2a, determiné que el estimador para la función de riesgo cuadrática (la media) es 2.0588, el estimador para la función de riesgo discreta (el modo) es 2.0392, y el estimador para la función de riesgo lineal (la mediana) es 2.0523. Estos valores indican que, bajo diferentes criterios de optimización, el número más representativo de llamadas al call center en un intervalo de dos minutos es ligeramente superior a dos, proporcionando una base sólida para decisiones operativas y de planificación.

Análisis de la Actividad 2c: Intervalo de Credibilidad Bayesiano

Finalmente, en la Actividad 2c, calculé un intervalo de credibilidad del 95% para el parámetro ϑ, utilizando la distribución posterior Gamma. Este intervalo, que va de 1.6839 a 2.4709, ofrece una medida de incertidumbre y confianza en nuestras estimaciones de la tasa de llamadas. La interpretación práctica de este resultado es que hay un 95% de probabilidad de que la tasa verdadera de llamadas al call center esté entre aproximadamente 1.68 y 2.47 llamadas por intervalo de dos minutos. Este intervalo de credibilidad proporciona una herramienta valiosa para evaluar la variabilidad y asegurar la adecuada asignación de recursos en respuesta a la demanda prevista.

```
C:\Users\matia\Documents\Matias1\Inferencia bayesiana\revaluacion>python actividad2.py
Parámetros de la distribución Gamma posterior: alpha=105, beta=51
Estimador bayesiano para la función de riesgo cuadrática (media): 2.0588235294117645
Estimador bayesiano para la función de riesgo discreta (modo): 2.0392156862745097
Estimador bayesiano para la función de riesgo lineal (mediana): 2.052291283253961
Intervalo bayesiano de credibilidad del 95%: (1.6839136729466304, 2.4708511175491243)
```

Matias Davila
Inferencia Bayesiana

Análisis de la Actividad 3:

Para la Actividad 3, analizamos datos de visitas a un medio de comunicación para evaluar el interés en noticias sobre conflictos bélicos. Al realizar el contraste bayesiano de hipótesis, calculamos las probabilidades posteriores de dos supuestos valores de ϑ, que representan la tasa promedio de visitas diarias.

La hipótesis $H0$ sostiene que la tasa promedio de visitas es de 330 por día, mientras que $H1$ postula que es de 480. Utilizando una distribución previa no informativa, que asigna igual credibilidad a ambas hipótesis antes de observar los datos, aplicamos el teorema de Bayes para actualizar estas creencias.

Las probabilidades posteriores calculadas reflejan la fuerza de la evidencia en favor de cada hipótesis. Con un valor de aproximadamente 0.656, la hipótesis $H1$ es favorecida sobre $H0$, que obtuvo una probabilidad de 0.344. Esto indica que los datos observados son más coherentes con una tasa promedio más alta de visitas, específicamente 480 visitas por día.

Esta información es crítica para el medio de comunicación al planificar contenido y recursos. Reconociendo que el interés en noticias de conflictos bélicos podría ser más alto de lo inicialmente asumido, la empresa podría asignar más recursos a esta área para satisfacer la demanda de los lectores y, potencialmente, atraer una audiencia más amplia.

En Conclusión, la Actividad 3 demuestra la aplicación efectiva de la inferencia bayesiana en el ámbito de la toma de decisiones basada en datos. Con un enfoque que va más allá de la simple estimación de parámetros, hemos podido cuantificar la incertidumbre y proporcionar estimaciones más informativas que pueden influir directamente en la estrategia de contenidos del medio de comunicación.

```
C:\Users\matia\Documents\Matias1\Inferencia bayesiana\revaluacion>python actividad3.py
Probabilidad posterior para H0 (theta = 330): 0.3436061248216544
Probabilidad posterior para H1 (theta = 480): 0.6563938751783289
```

Matias Davila
Inferencia Bayesiana

Análisis de la Actividad 4: Distribución Predictiva Posterior para y^*=0

En la Actividad 4, mi objetivo fue calcular la distribución predictiva posterior para una observación futura y^*=0 en un contexto de modelado Bernoulli-Beta. Asumiendo una distribución previa con una cierta cantidad de éxitos y fracasos y los parámetros previos de Beta α=2 y β=2, actualicé los parámetros basándome en los datos observados.

Los parámetros de la distribución posterior Beta fueron actualizados a α'=12 y β'=7 después de contabilizar 10 éxitos y 5 fracasos en los datos. Utilizando esta información actualizada, la probabilidad predictiva posterior para la observación futura de y^*=0 fue calculada como aproximadamente 0.3684.

Este resultado significa que, dada la evidencia acumulada y las creencias previas ajustadas, hay una probabilidad de alrededor de 36.84% de que la próxima observación en esta serie de datos sea un fracaso (y^*=0). Esta estimación es crucial para entender la tendencia en los resultados y puede servir como una guía para las expectativas futuras en situaciones similares.

La inferencia bayesiana nos ha permitido no solo actualizar nuestra creencia sobre el parámetro de interés sino también hacer predicciones sobre eventos futuros. Este enfoque es fundamentalmente poderoso en la toma de decisiones basada en datos, ya que nos permite cuantificar la incertidumbre y ajustar nuestras predicciones a medida que se recopilan más datos.

```
C:\Users\matia\Documents\Matias1\Inferencia bayesiana\revaluacion>python actividad4.py
Probabilidad predictiva posterior para y* = 0: 0.3684210526315789
```

Matias Davila
Inferencia Bayesiana

UNIVERSITAT
Carlemany

Análisis de la Actividad 5A: Caracterización de la Distribución Posterior

En la Actividad 5A, utilicé una distribución Beta previa que reflejaba la creencia de la alta calidad de los coches, según lo afirmado por el fabricante. Sin embargo, tras la observación de que 12 de 20 coches comprados necesitaron reparación durante el periodo de garantía, actualicé la distribución previa Beta(4.75, 0.25) con estos nuevos datos. Los resultados son indicativos del rendimiento real de los coches y proporcionan una visión bayesiana más informada que la creencia inicial.

Los cálculos resultaron en una **media posterior de 0.67**, lo que indica que, en promedio, la probabilidad de que un coche requiera reparación es del 67% en contraposición al 95% de confiabilidad anunciado por el fabricante. La **mediana posterior de 0.675** es cercana a la media, lo que sugiere una distribución simétrica y reafirma la estimación central de la tasa de fallos.

El **intervalo de credibilidad del 95% de (0.478, 0.836)** nos dice que, con un alto nivel de confianza, la verdadera tasa de fallos de los coches se encuentra entre aproximadamente 47.8% y 83.6%. Este intervalo amplio refleja una incertidumbre significativa en la tasa de fallos, pero claramente indica que la fiabilidad de los coches es menor de lo que se esperaba inicialmente.

Este análisis sugiere que la empresa de transportes podría necesitar reconsiderar su evaluación de la fiabilidad de los coches basándose en la evidencia empírica recogida, lo cual difiere notablemente de las afirmaciones del fabricante. La actualización de creencias a través del enfoque bayesiano ha proporcionado una nueva perspectiva sobre la calidad de los vehículos y puede guiar decisiones futuras sobre la adquisición de coches o la negociación de términos de garantía.

```
C:\Users\matia\Documents\Matias1\Inferencia bayesiana\revaluacion>python actividad5.py
Media posterior: 0.67
Mediana posterior: 0.6745984749883508
Intervalo de credibilidad del 95%: (0.47834390762838225, 0.835913636746955)
```

Análisis de los Resultados de la Simulación Monte Carlo:

1. **Resultados con tamaño de muestra de 30**:
    - **Media**: 0.681
    - **Mediana**: 0.678
    - **Intervalo de credibilidad del 95%**: [0.518, 0.820]

Con el tamaño de muestra más pequeño, observamos una mayor variabilidad en el intervalo de credibilidad. Esto es típico debido a la menor cantidad de datos que pueden captar menos la variabilidad real de la distribución posterior.

2. **Resultados con tamaño de muestra de 3000**:
   - **Media**: 0.673
   - **Mediana**: 0.678
   - **Intervalo de credibilidad del 95%**: [0.481, 0.831]

Aumentando el tamaño de muestra, la estimación se vuelve más precisa y el intervalo de credibilidad se estabiliza, mostrando menos variabilidad y una mejor captura de la verdadera incertidumbre de la distribución posterior.

3. **Resultados con tamaño de muestra de 30000**:
   - **Media**: 0.669
   - **Mediana**: 0.674
   - **Intervalo de credibilidad del 95%**: [0.475, 0.834]

Con el tamaño de muestra más grande, las estimaciones se estabilizan aún más, acercándose a lo que sería una aproximación muy precisa de la verdadera distribución posterior. Esto indica que las simulaciones de Monte Carlo con grandes tamaños de muestra pueden proporcionar estimaciones muy fiables y consistentes.

Conclusiones del Análisis:

Las simulaciones de Monte Carlo han demostrado ser una herramienta poderosa para entender la distribución posterior de la probabilidad de falla de los coches. A medida que aumentamos el tamaño de la muestra, las estimaciones se vuelven más precisas y el intervalo de credibilidad se hace más estrecho, lo que proporciona una mayor confianza en nuestros resultados.

Este enfoque nos permite cuantificar la incertidumbre de manera efectiva y ajustar nuestras estimaciones basadas en una mayor cantidad de simulaciones, lo cual es crucial para decisiones informadas en el contexto de gestión de garantías y evaluación de la calidad del producto.

```
C:\Users\matia\Documents\Matias1\Inferencia bayesiana\revaluacion>python actividad5b.py

Tamaño de muestra: 30
Media estimada por Monte Carlo: 0.6806203153192553
Mediana estimada por Monte Carlo: 0.6778877021987595
Intervalo de credibilidad del 95% estimado por Monte Carlo: [0.51833326 0.81995258]

Tamaño de muestra: 3000
Media estimada por Monte Carlo: 0.6727252163140156
Mediana estimada por Monte Carlo: 0.677536225973244
Intervalo de credibilidad del 95% estimado por Monte Carlo: [0.4813394  0.83056031]

Tamaño de muestra: 30000
Media estimada por Monte Carlo: 0.6691285352216362
Mediana estimada por Monte Carlo: 0.6741484357031889
Intervalo de credibilidad del 95% estimado por Monte Carlo: [0.47544766 0.83442856]
```