Análisis de Sistemas Dinámicos y Estimación
(EL3204-1)
Clase auxiliar 7
Prof. Marcos Orchard - Sebastián Espinosa.
Prof. Aux. Erik Sáez

1 Resumen

Variables Aleatorias

Una **variable aleatoria** X es una función que asigna un número real a cada resultado de un experimento aleatorio: $X : \Omega \longrightarrow \mathbb{R}$. En la práctica, trabajamos con las distribuciones de estas variables.

Las variables continuas tienen una **función de densidad de probabilidad** $f(x)$ que describe qué tan "densa" es la probabilidad alrededor de cada punto. No da probabilidades directas, sino que las probabilidades se calculan integrando: $\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)\,dx$. Debe cumplir $f(x) \geq 0$ y $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

Las variables discretas tienen una **función de probabilidad de masa** $p(x)$ que da directamente la probabilidad de cada valor específico. Debe cumplir $p(x) \geq 0$ y $\sum_x p(x) = 1$. Las probabilidades se calculan sumando: $\mathbb{P}(X \in A) = \sum_{x \in A} p(x)$.

Distribuciones Importantes

Distribución Uniforme $X \sim U(a, b)$: Todos los valores en $[a, b]$ son igualmente probables.

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{en otro caso} \end{cases} \tag{1}$$

Distribución Exponencial $X \sim \text{Exp}(\lambda)$: Modela tiempos de espera entre eventos.

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{en otro caso} \end{cases} \tag{2}$$

Distribución Normal $X \sim \mathcal{N}(\mu, \sigma^2)$: La distribución más importante. Curva en forma de campana.

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{3}$$

Propiedad clave: cualquier combinación lineal de variables normales independientes también es normal.

Distribución Poisson $X \sim \text{Poisson}(\lambda)$: Modela el número de eventos raros en un intervalo fijo.

$$p(x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!} & \text{si } x \in \{0, 1, 2, \ldots\} \\ 0 & \text{en otro caso} \end{cases} \tag{4}$$

Vectores Aleatorios

Un **vector aleatorio** $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ agrupa múltiples variables aleatorias. Su **función de densidad conjunta** describe el comportamiento conjunto:

$$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) \tag{5}$$

Para variables **independientes e idénticamente distribuidas (i.i.d.)**:

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^{n} f_{X_i}(x_i) \tag{6}$$

Momentos Estadísticos

Los **momentos** caracterizan las propiedades de una distribución:
Esperanza (valor promedio esperado):

$$\mathbb{E}[X] = \begin{cases} \int_{-\infty}^{\infty} x \cdot f(x)\,dx & \text{si } X \text{ es continua} \\ \sum_x x \cdot p(x) & \text{si } X \text{ es discreta} \end{cases} \tag{7}$$

Varianza (medida de dispersión):

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \tag{8}$$

Covarianza (dependencia lineal entre dos variables):

$$\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \tag{9}$$

Para la distribución normal $X \sim \mathcal{N}(\mu, \sigma^2)$: $\mu = \mathbb{E}[X]$ y $\sigma^2 = \mathrm{Var}(X)$.

1. Responda las siguientes preguntas:

   1. Sea X una variable aleatoria discreta tal que $X \sim \text{poisson}(\lambda)$, obtenga su esperanza y su varianza.
   2. Sean $X_1, X_2, \ldots, X_n$ variables aleatorias continuas i.i.d tal que cada una distribuye $\mathcal{N}(\mu, \sigma^2)$. Obtenga su distribución conjunta.
   3. Sean $X_1, X_2, \ldots, X_n$ variables aleatorias continuas i.i.d tal que cada una distribuye $\mathcal{N}(\mu, \sigma^2)$. Obtenga la distribución de $Y = \sum_{i=1}^{n} \frac{X_i - \mu}{\sqrt{n}\sigma}$.

Solución:

Resolución 1.1

Para una variable aleatoria X que sigue una distribución de Poisson con parámetro λ, recordemos que su función de probabilidad de masa está dada por $P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$ para $x = 0, 1, 2, \ldots$

Para obtener la esperanza, aplicamos la definición correspondiente a variables aleatorias discretas. Calculamos $\mathbb{E}[X] = \sum_{x=0}^{\infty} x \cdot P(X = x) = \sum_{x=0}^{\infty} x \cdot \frac{e^{-\lambda}\lambda^x}{x!}$. Observamos que el término con $x = 0$ es igual a cero, por lo que podemos reescribir la suma como:

$$\mathbb{E}[X] = \sum_{x=1}^{\infty} x \cdot \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^x}{(x-1)!} = e^{-\lambda}\lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \tag{10}$$

Al realizar el cambio de variable $u = x - 1$, obtenemos:

$$\mathbb{E}[X] = e^{-\lambda}\lambda \sum_{u=0}^{\infty} \frac{\lambda^u}{u!} \tag{11}$$

La sumatoria que aparece corresponde exactamente a la serie de Taylor de la función exponencial $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$. Por tanto, $\sum_{u=0}^{\infty} \frac{\lambda^u}{u!} = e^{\lambda}$, lo que nos da:

$$\mathbb{E}[X] = e^{-\lambda} \cdot \lambda \cdot e^{\lambda} = \lambda \tag{12}$$

Para calcular la varianza, utilizamos la fórmula $\text{Var}(X) = \mathbb{E}[X^2] - [\mathbb{E}[X]]^2$ y necesitamos determinar $\mathbb{E}[X^2]$. Siguiendo un procedimiento similar al anterior:

$$\mathbb{E}[X^2] = \sum_{x=0}^{\infty} x^2 \cdot \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=1}^{\infty} x^2 \cdot \frac{e^{-\lambda}\lambda^x}{x!} \tag{13}$$

$$= \sum_{x=1}^{\infty} x \cdot \frac{e^{-\lambda}\lambda^x}{(x-1)!} = e^{-\lambda}\lambda \sum_{x=1}^{\infty} x \cdot \frac{\lambda^{x-1}}{(x-1)!} \tag{14}$$

Con el cambio de variable $u = x - 1$ (de modo que $x = u + 1$), ademas recordemos que $\sum_{u=0}^{\infty} \frac{\lambda^u}{u!} = e^{\lambda}$ y $\sum_{u=0}^{\infty} u \cdot \frac{\lambda^u}{u!} = \lambda e^{\lambda}$, tenemos:

$$\mathbb{E}[X^2] = e^{-\lambda}\lambda \sum_{u=0}^{\infty} (u+1) \cdot \frac{\lambda^u}{u!} \tag{15}$$

$$= e^{-\lambda}\lambda \left[\sum_{u=0}^{\infty} u \cdot \frac{\lambda^u}{u!} + \sum_{u=0}^{\infty} \frac{\lambda^u}{u!} \right] \tag{16}$$

$$= e^{-\lambda}\lambda \left[\lambda e^{\lambda} + e^{\lambda} \right] = \lambda^2 + \lambda \tag{17}$$

Finalmente, la varianza resulta ser:

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - [\mathbb{E}[X]]^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda \tag{18}$$

Concluimos que para una variable aleatoria $X \sim \mathrm{Poisson}(\lambda)$, tanto su esperanza como su varianza son iguales al parámetro λ.

Resolución 1.2

Necesitamos encontrar la función de densidad conjunta de n variables aleatorias $X_1, X_2, \ldots, X_n$ que son independientes e idénticamente distribuidas con distribución $\mathcal{N}(\mu, \sigma^2)$.

Cuando las variables aleatorias son independientes, una propiedad fundamental nos dice que la función de densidad conjunta es simplemente el producto de las funciones de densidad marginales. Es decir:

$$f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i) \tag{19}$$

Sabemos que cada variable aleatoria individual $X_i \sim \mathcal{N}(\mu, \sigma^2)$ tiene función de densidad:

$$f_{X_i}(x_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right) \tag{20}$$

Al aplicar la propiedad de independencia y multiplicar todas las densidades marginales:

$$\begin{aligned} f_{X_1,X_2,\ldots,X_n}(x_1, \ldots, x_n) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right) && (21) \\ &= \frac{1}{(\sqrt{2\pi}\sigma)^n} \prod_{i=1}^{n} \exp\left(-\frac{(x_i - \mu)^2}{2\sigma^2}\right) && (22) \\ &= \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\right) && (23) \end{aligned}$$

Por tanto, la función de densidad conjunta está dada por:

$$f_{X_1,X_2,\ldots,X_n}(x_1, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\right) \tag{24}$$

Esta expresión corresponde a una **distribución normal multivariada** (o gaussiana multivariada), que generalizamos como $\mathbf{X} = (X_1, \ldots, X_n) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, donde:

- **Vector de medias:** $\boldsymbol{\mu} = (\mu, \mu, \ldots, \mu)^T \in \mathbb{R}^n$

  Cada componente del vector aleatorio tiene la misma media μ, es decir, $\mathbb{E}[X_i] = \mu$ para todo $i = 1, \ldots, n$.

- **Matriz de covarianza:** $\mathbf{\Sigma} = \sigma^2 I_n$, donde I_n es la matriz identidad de orden n:

$$\mathbf{\Sigma} = \sigma^2 \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{pmatrix} \tag{25}$$

  Los elementos de la diagonal principal son $\mathrm{Var}(X_i) = \sigma^2$ para todo i, mientras que los elementos fuera de la diagonal son $\mathrm{Cov}(X_i, X_j) = 0$ para $i \neq j$.

- **Independencia:** La forma diagonal de $\mathbf{\Sigma}$ (con ceros fuera de la diagonal) refleja precisamente que las variables son independientes. En general, para variables normales, covarianza cero implica independencia estadística. Por tanto:

$$\mathrm{Cov}(X_i, X_j) = 0 \text{ para } i \neq j \quad \Longleftrightarrow \quad X_i \text{ y } X_j \text{ son independientes} \tag{26}$$

En resumen, podemos escribir que el vector aleatorio $\mathbf{X} = (X_1, \ldots, X_n) \sim \mathcal{N}(\boldsymbol{\mu}, \sigma^2 I_n)$, donde cada variable tiene la misma media μ y varianza σ^2, y todas son mutuamente independientes.

Resolución 1.3

Se nos pide encontrar la distribución de $Y = \sum_{i=1}^{n} \frac{X_i - \mu}{\sqrt{n}\sigma}$, donde $X_1, X_2, \ldots, X_n$ son variables aleatorias i.i.d. con distribución $\mathcal{N}(\mu, \sigma^2)$.

Como la variable Y es una combinación lineal de variables aleatorias normales independientes, también distribuirá normal (propiedad vista en el recordatorio). Con ello, calculamos la esperanza y su varianza para tener así sus parámetros. Para la esperanza, se tiene:

$$\begin{aligned} \mathbb{E}[Y] &= \mathbb{E}\left[\sum_{i=1}^{n} \frac{X_i - \mu}{\sqrt{n}\sigma}\right] = \frac{1}{\sqrt{n}\sigma}\mathbb{E}\left[\sum_{i=1}^{n}(X_i - \mu)\right] && (27) \\ &= \frac{1}{\sqrt{n}\sigma}\left[\sum_{i=1}^{n}\mathbb{E}[X_i - \mu]\right] = \frac{1}{\sqrt{n}\sigma}\left[\sum_{i=1}^{n}(\mathbb{E}[X_i] - \mu)\right] && (28) \\ &= \frac{1}{\sqrt{n}\sigma}\left[\sum_{i=1}^{n}(\mu - \mu)\right] = \frac{1}{\sqrt{n}\sigma} \cdot 0 = 0 && (29) \end{aligned}$$

Para lo cual se utilizo la propiedad de la esperanza dada por:

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b \tag{30}$$

Para la varianza, utilizaremos que son i.i.d., pues si X y Y son dos v.a. independientes se cumple que $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$:

$$\mathrm{Var}(Y) = \mathrm{Var}\left(\sum_{i=1}^{n} \frac{X_i - \mu}{\sqrt{n}\sigma}\right) = \sum_{i=1}^{n} \mathrm{Var}\left(\frac{X_i - \mu}{\sqrt{n}\sigma}\right) \tag{31}$$

Donde se utiliza la propiedad de la varianza dada por:

$$\mathrm{Var}(aX+b) = a^2\mathrm{Var}(X) \tag{32}$$

Continuando con el cálculo de la varianza:

$$= \sum_{i=1}^{n} \frac{1}{n\sigma^2}\mathrm{Var}(X_i - \mu) = \sum_{i=1}^{n} \frac{1}{n\sigma^2}\mathrm{Var}(X_i) = \sum_{i=1}^{n} \frac{1}{n\sigma^2} \cdot \sigma^2 \tag{33}$$

$$= \sum_{i=1}^{n} \frac{1}{n} = \frac{n \cdot 1}{n} = 1. \tag{34}$$

Finalmente, tenemos que $Y \sim \mathcal{N}(0,1)$, lo cual significa que se acaba de realizar una estandarización (Le restamos a una variable su media y dividimos la resta por su desviación estándar).

2. Suponga que usted intenta comprobar la robustez de un canal de comunicaciones. Para ello, se contacta con su amigo para encomendarle que se coloque en el extremo emisor del canal y que decida de entre las señales $\{0,1\}$ con el lanzamiento de una moneda equilibrada (siendo así aleatorio). Mientras tanto, usted se encuentra en el extremo receptor del canal con un sensor perfecto. Definiendo θ como la señal y $N \sim \mathcal{N}(0, \sigma_0^2)$ el ruido gaussiano debido al canal, se tiene:

$$Y = \theta + N \tag{35}$$

Siendo Y por lo tanto, la medición de la señal en el extremo receptor. El fabricante del canal le informa que la mejor cifra para determinar entre ambas señales es un valor τ, de tal manera que si $Y < \tau$ usted dirá que es la señal 0, en caso contrario dirá 1.

1. Calcule la probabilidad de que dado que $\theta = 0$ usted diga que la señal sea $\theta = 0$. Interprete que significa esto.
2. Calcule la probabilidad de que dado que $\theta = 1$ usted diga que la señal sea $\theta = 0$. Interprete que significa esto.
3. Calcule la probabilidad de que dado que $\theta = 1$ usted diga que la señal sea $\theta = 1$. Interprete que significa esto.
4. Calcule la probabilidad de que dado que $\theta = 0$ usted diga que la señal sea $\theta = 1$. Interprete que significa esto.
5. Obtenga la probabilidad de error.
   Hint: Le puede ser útil el teorema de probabilidades totales.
   Sean A y B una partición de todo el espacio muestral Ω y un evento C, entonces se cumple:

$$\mathbb{P}(C) = \mathbb{P}(C|A) \cdot \mathbb{P}(A) + \mathbb{P}(C|B) \cdot \mathbb{P}(B) \tag{36}$$

6. Obtenga la probabilidad de acertar. *Hint:* use el mismo teorema que en c).
7. ¿Cómo **minimizaría** la probabilidad de error?
8. ¿Cómo **maximizaría** la probabilidad de acertar?

Solución:

Resolución 2.1

Necesitamos calcular la probabilidad de que dado que $\theta = 0$ usted diga que la señal sea $\theta = 0$. Primero notemos que la probabilidad pedida es que dado que $\theta = 0$, usted diga que $\theta = 0$, es decir, que $Y \leq \tau$. Luego, pedimos $\mathbb{P}(Y \leq \tau | \theta = 0)$.

Si $\theta = 0$, entonces $Y = N$ y se cumple que $Y|\theta = 0 \sim \mathcal{N}(0, \sigma_0^2)$. Por tanto, la función de densidad es:

$$f_{Y|\theta=0} = \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{y^2}{2\sigma_0^2}} \tag{37}$$

Con ello, la probabilidad pedida corresponde a:

$$\mathbb{P}(Y \leq \tau | \theta = 0) = \int_{-\infty}^{\tau} \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{y^2}{2\sigma_0^2}} \, dy \tag{38}$$

Esta probabilidad corresponde a detectar correctamente la ausencia de señal.

Resolución 2.2

La probabilidad pedida en este caso corresponde a que dado que $\theta = 1$ y si se decide una señal, el valor de Y debe ser menor a τ y en consecuencia usted diga que es 0. Si $\theta = 1$ entonces $Y = 1 + N$ y por lo tanto $Y|\theta = 1 \sim \mathcal{N}(1, \sigma_0^2)$. La función de densidad es:

$$f_{Y|\theta=1} = \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{(y-1)^2}{2\sigma_0^2}} \tag{39}$$

Donde recordemos la propiedad de que al sumar una constante a una variable aleatoria normal, la media se desplaza en esa constante y la varianza permanece igual, es decir:

$$\text{Si } X \sim \mathcal{N}(\mu, \sigma^2) \text{ entonces } X + c \sim \mathcal{N}(\mu + c, \sigma^2) \tag{40}$$

Con ello, la probabilidad pedida corresponde a:

$$\mathbb{P}(Y \leq \tau | \theta = 1) = \int_{-\infty}^{\tau} \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{(y-1)^2}{2\sigma_0^2}} \, dy \tag{41}$$

La probabilidad pedida corresponde a equivocarse en la detección de la señal cuando esta realmente es emitida. En jerga de detección es conocido como error tipo II. Este tipo error puede ser catastrófico dependiendo del contexto. Por ejemplo, si el evento a detectar es una enfermedad, el error tipo II corresponde a que el paciente tenga la enfermedad, pero el diagnóstico realizado no la detecte.

Resolución 2.3

Análogo al caso anterior, la probabilidad pedida es $\mathbb{P}(Y \geq \tau|\theta = 1)$. Con lo cual la probabilidad queda en la siguiente expresión:

$$\mathbb{P}(Y \geq \tau|\theta = 1) = \int_{\tau}^{\infty} \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{(y-1)^2}{2\sigma_0^2}} \, dy \tag{42}$$

La probabilidad pedida corresponde a detectar correctamente la aparición de la señal. En jerga de detección de test, este valor se suele llamar potencia del test. Análogo al caso anterior, la probabilidad pedida es $\mathbb{P}(Y \geq \tau|\theta = 1)$. Con lo cual la probabilidad queda en la siguiente expresión:

$$\mathbb{P}(Y \geq \tau|\theta = 1) = \int_{\tau}^{\infty} \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{(y-1)^2}{2\sigma_0^2}} \, dy \tag{43}$$

Resolución 2.4

En este caso, ocurre que $\theta = 0$ y queremos saber la probabilidad de que se diga que hay señal ($\theta = 1$), es decir, que $Y \geq \tau$. Luego, ocuparemos la densidad $f_{Y|\theta=0}$. Con esto, la probabilidad pedida queda en la siguiente expresión:

$$\mathbb{P}(Y \geq \tau|\theta = 0) = \int_{\tau}^{\infty} \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{y^2}{2\sigma_0^2}} \, dy \tag{44}$$

La probabilidad pedida corresponde a equivocarse al dar una falsa detección de la señal cuando realmente no fue emitida. En jerga de detección es conocido como error tipo I o falsa alarma.

Resolución 2.5

La probabilidad de error no puede ser obtenida directamente debido a que tenemos dos fuentes posibles de error, que serían:

- Dado que la señal es $\theta = 0$, decir $\theta = 1$.
- Dado que la señal es $\theta = 1$, decir $\theta = 0$.

Luego, si tomamos la partición de los dos valores posibles de theta y usamos el teorema de probabilidades totales mostrado como hint, podemos obtener la probabilidad de error:

$$\mathbb{P}(error) = \mathbb{P}(error|\theta = 0) \cdot \mathbb{P}(\theta = 0) + \mathbb{P}(error|\theta = 1) \cdot \mathbb{P}(\theta = 1) \tag{45}$$

Como las señales son enviadas con igual probabilidad (debido a que se decide cuál enviar con el lanzamiento de una moneda equilibrada), tenemos que $\mathbb{P}(\theta = 0) = \mathbb{P}(\theta = 1) = \frac{1}{2}$. Respecto a las otras probabilidades, se puede notar que $\mathbb{P}(error|\theta = 0) = \mathbb{P}(Y \geq \tau|\theta = 0)$ y $\mathbb{P}(error|\theta = 1) = \mathbb{P}(Y \leq \tau|\theta = 1)$. Luego, reemplazando los valores usando lo calculado en los ítems anteriores, tenemos:

$$\mathbb{P}(error) = \frac{1}{2} \cdot \int_{\tau}^{\infty} \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{y^2}{2\sigma_0^2}} \, dy + \frac{1}{2} \int_{-\infty}^{\tau} \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{(y-1)^2}{2\sigma_0^2}} \, dy \tag{46}$$

Observación: El uso del teorema de probabilidades totales con las probabilidades a priori $\mathbb{P}(\theta = 0) = \mathbb{P}(\theta = 1) = \frac{1}{2}$ indica que estamos en un contexto bayesiano, donde tratamos a θ como una variable aleatoria con distribución conocida.

Resolución 2.6

El cálculo de la probabilidad de acertar es similar al cálculo de probabilidad de error. Se tienen dos casos posibles para acertar:

- Dado que la señal es $\theta = 0$, decir $\theta = 0$.
- Dado que la señal es $\theta = 1$, decir $\theta = 1$.

Luego, usamos la misma partición que en c) y reemplazamos:

$$\mathbb{P}(acertar) = \mathbb{P}(acertar|\theta = 0) \cdot \mathbb{P}(\theta = 0) + \mathbb{P}(acertar|\theta = 1) \cdot \mathbb{P}(\theta = 1) \tag{47}$$

Se puede notar que $\mathbb{P}(acertar|\theta = 0) = \mathbb{P}(Y \leq \tau|\theta = 0)$ y $\mathbb{P}(acertar|\theta = 1) = \mathbb{P}(Y \geq \tau|\theta = 1)$. Luego, reemplazando en la ecuación:

$$\mathbb{P}(acertar) = \frac{1}{2} \cdot \int_{-\infty}^{\tau} \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{y^2}{2\sigma_0^2}} \, dy + \frac{1}{2} \int_{\tau}^{\infty} \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{(y-1)^2}{2\sigma_0^2}} \, dy \tag{48}$$

Resolución 2.7

Para minimizar la probabilidad de error, se debe notar que el único parámetro que puede variar en el problema es el valor τ, ya que su modificación puede afectar cuántas detecciones realizamos correctamente. En base a esto, podemos definir la función probabilidad de error en función de τ como:

$$e(\tau) = \frac{1}{2} \int_{\tau}^{\infty} \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{y^2}{2\sigma_0^2}} \, dy + \frac{1}{2} \int_{-\infty}^{\tau} \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{(y-1)^2}{2\sigma_0^2}} \, dy \tag{49}$$

Para minimizar la probabilidad de error, usaremos el criterio de la primera derivada tal que $\frac{de}{d\tau} = 0$. Luego:

$$\frac{d}{d\tau} \left(\frac{1}{2} \int_{\tau}^{\infty} \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{y^2}{2\sigma_0^2}} \, dy + \frac{1}{2} \int_{-\infty}^{\tau} \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{(y-1)^2}{2\sigma_0^2}} \, dy \right) = 0 \tag{50}$$

Utilizando el Teorema Fundamental del Cálculo, que establece que $\frac{d}{dx} \int_a^x f(u)du = f(x)$, llegamos a:

$$\begin{aligned}
-\frac{1}{2} \cdot \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{\tau^2}{2\sigma_0^2}} + \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{(\tau-1)^2}{2\sigma_0^2}} &= 0 && (51)\\
\frac{1}{2} \cdot \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{\tau^2}{2\sigma_0^2}} &= \frac{1}{2} \cdot \frac{1}{\sqrt{2\pi} \cdot \sigma_0} \cdot e^{-\frac{(\tau-1)^2}{2\sigma_0^2}} && (52)\\
e^{-\frac{\tau^2}{2\sigma_0^2}} &= e^{-\frac{(\tau-1)^2}{2\sigma_0^2}} && (53)\\
\frac{-\tau^2}{2\sigma_0^2} &= \frac{-(\tau-1)^2}{2\sigma_0^2} && (54)\\
\tau^2 &= (\tau-1)^2 && (55)\\
\tau^2 &= \tau^2 - 2\tau + 1 && (56)\\
0 &= -2\tau + 1 && (57)\\
\tau &= \frac{1}{2} && (58)
\end{aligned}$$

Luego, el valor que minimiza el valor de la probabilidad de error es $\tau = \frac{1}{2}$.

Resolución 2.8

El cálculo es análogo al ítem anterior, y se llega al valor $\tau = \frac{1}{2}$. Lo interesante es notar que dicho valor genera que la probabilidad de error se minimice y que la probabilidad de acertar se maximice. Como dato curioso, esto corresponde al punto medio entre las medias de ambas funciones y el punto de intersección de ambas funciones de densidad condicionales.

Nota importante: Es importante señalar que en este problema hemos tratado a θ como un parámetro fijo (enfoque frecuentista o paramétrico). Sin embargo, dado que el enunciado menciona que la señal se decide "con el lanzamiento de una moneda equilibrada", técnicamente θ es una variable aleatoria, lo que situaría el problema en un contexto bayesiano. Este enfoque bayesiano, donde las hipótesis tienen probabilidades a priori, será tratado con mayor profundidad en clases posteriores del curso.

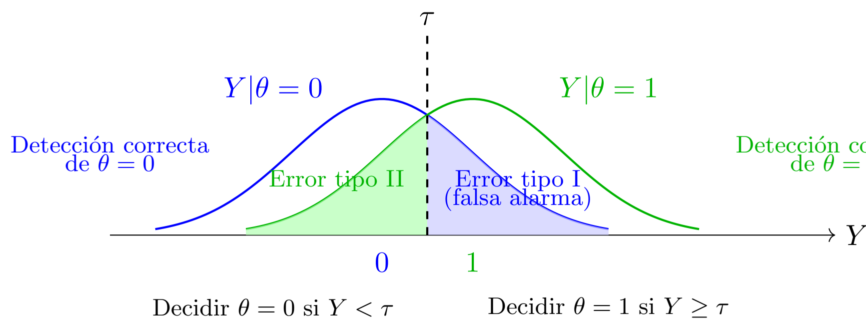

Figura 1: Distribuciones condicionales del problema de detección de señal: $Y|\theta = 0 \sim \mathcal{N}(0, \sigma_0^2)$ y $Y|\theta = 1 \sim \mathcal{N}(1, \sigma_0^2)$ con umbral de decisión τ. Las regiones sombreadas ilustran los errores tipo I (falsa alarma) y tipo II.

3. Una máquina de acuñación defectuosa produce monedas cuya probabilidad de obtener cara es una variable aleatoria P con PDF

$$f_P(p) = \begin{cases} 1 + \sin(2\pi p), & \text{si } p \in [0,1], \\ 0, & \text{en otro caso.} \end{cases} \tag{59}$$

En esencia, una moneda específica producida por esta máquina tendrá una probabilidad fija $P = p$ de dar cara, pero usted no conoce inicialmente cuál es esa probabilidad. Una moneda producida por esta máquina es seleccionada y lanzada repetidamente, con lanzamientos sucesivos asumidos independientes.

1. Encuentre la probabilidad de que el primer lanzamiento de la moneda resulte en cara.
2. Dado que el primer lanzamiento de la moneda resultó en cara, encuentre la PDF condicional de P.
3. Dado que el primer lanzamiento de la moneda resultó en cara, encuentre la probabilidad condicional de obtener cara en el segundo lanzamiento.

Solución:

Resolución 3.1

Primero, es crucial entender qué representa cada elemento del problema:

- **La variable P:** No es un número fijo, sino una *variable aleatoria* que representa la probabilidad de obtener cara de una moneda específica producida por la máquina defectuosa. Cada moneda que produce la máquina tiene su propia probabilidad P de dar cara, pero esta probabilidad varía de moneda a moneda.
- **La distribución $f_P(p) = 1+\sin(2\pi p)$:** Esta función nos dice qué tan probable es que una moneda producida por la máquina tenga probabilidad p de dar cara. Es una distribución que:
  - Oscila entre $1 - 1 = 0$ y $1 + 1 = 2$
  - Tiene máximos en $p = 0.25$ y $p = 0.75$ (donde $\sin(2\pi p) = 1$)
  - Tiene mínimos en $p = 0$ y $p = 0.5$ y $p = 1$ (donde $\sin(2\pi p) = 0$ o -1)
  - Esto significa que la máquina tiende a producir monedas sesgadas hacia $p = 0.25$ o $p = 0.75$, pero raramente produce monedas justas ($p = 0.5$)
- **El proceso:** Tomamos una moneda al azar de las producidas por esta máquina defectuosa. No sabemos cuál es su probabilidad específica P, solo sabemos que sigue la distribución dada.

El problema fundamental aquí es que no conocemos el valor exacto de la probabilidad P de la moneda específica que tenemos. Esto significa que debemos considerar todos los posibles valores que P puede tomar, ponderando cada uno por su probabilidad de ocurrencia según la distribución $f_P(p)$ que nos proporciona el enunciado.

Para calcular la probabilidad $\mathbb{P}(A)$, usamos la versión continua del teorema de probabilidad total:

$$\mathbb{P}(A) = \int_0^1 \mathbb{P}(A|P = p) f_P(p)\, dp = \int_0^1 p(1 + \sin(2\pi p))\, dp, \tag{60}$$

donde $\mathbb{P}(A|P = p) = p$ porque si la moneda tiene probabilidad fija p de dar cara, entonces la probabilidad de obtener cara es exactamente p. Separamos la integral en dos partes:

$$\mathbb{P}(A) = \int_0^1 p\, dp + \int_0^1 p \sin(2\pi p)\, dp \tag{61}$$

La primera integral es directa:

$$\int_0^1 p\, dp = \left[\frac{p^2}{2}\right]_0^1 = \frac{1^2}{2} - \frac{0^2}{2} = \frac{1}{2} \tag{62}$$

Para la segunda integral, $\int_0^1 p \sin(2\pi p)\, dp$, usamos integración por partes con:

$$u = p \quad \Rightarrow \quad du = dp \tag{63}$$

$$dv = \sin(2\pi p)\, dp \quad \Rightarrow \quad v = -\frac{1}{2\pi}\cos(2\pi p) \tag{64}$$

Aplicando la fórmula de integración por partes $\int u\, dv = uv - \int v\, du$:

$$\int_0^1 p \sin(2\pi p)\, dp = \left[p \cdot \left(-\frac{1}{2\pi}\cos(2\pi p)\right)\right]_0^1 - \int_0^1 \left(-\frac{1}{2\pi}\cos(2\pi p)\right) dp \tag{65}$$

$$= \left[-\frac{p}{2\pi}\cos(2\pi p)\right]_0^1 + \frac{1}{2\pi}\int_0^1 \cos(2\pi p)\, dp \tag{66}$$

Evaluamos cada término por separado:

- **Primer término:**

$$\left[-\frac{p}{2\pi}\cos(2\pi p)\right]_0^1 = -\frac{1}{2\pi}\cos(2\pi\cdot 1) - \left(-\frac{0}{2\pi}\cos(2\pi\cdot 0)\right) \tag{67}$$
$$= -\frac{1}{2\pi}\cos(2\pi) - 0 \tag{68}$$
$$= -\frac{1}{2\pi}\cdot 1 = -\frac{1}{2\pi} \tag{69}$$

donde usamos que $\cos(2\pi) = 1$.

- **Segundo término:**

$$\frac{1}{2\pi}\int_0^1 \cos(2\pi p)\,dp = \frac{1}{2\pi}\left[\frac{1}{2\pi}\sin(2\pi p)\right]_0^1 \tag{70}$$
$$= \frac{1}{(2\pi)^2}[\sin(2\pi\cdot 1) - \sin(2\pi\cdot 0)] \tag{71}$$
$$= \frac{1}{4\pi^2}[0-0] = 0 \tag{72}$$

donde usamos que $\sin(2\pi) = \sin(0) = 0$.

Por tanto:

$$\int_0^1 p\sin(2\pi p)\,dp = -\frac{1}{2\pi} + 0 = -\frac{1}{2\pi} \tag{73}$$

Combinando ambas integrales:

$$\mathbb{P}(A) = \int_0^1 p\,dp + \int_0^1 p\sin(2\pi p)\,dp \tag{74}$$
$$= \frac{1}{2} + \left(-\frac{1}{2\pi}\right) \tag{75}$$
$$= \frac{1}{2} - \frac{1}{2\pi} \tag{76}$$
$$= \frac{\pi - 1}{2\pi} \approx 0.341 \tag{77}$$

El resultado $\mathbb{P}(A) = \frac{\pi-1}{2\pi} \approx 0.341 < 0.5$ tiene sentido intuitivo:

- Si todas las monedas fueran justas ($P = 0.5$ siempre), tendríamos $\mathbb{P}(A) = 0.5$
- Si la distribución fuera uniforme ($f_P(p) = 1$), también tendríamos $\mathbb{P}(A) = 0.5$
- Pero nuestra distribución $f_P(p) = 1 + \sin(2\pi p)$ favorece valores como $p = 0.25$ y $p = 0.75$
- El término $-\frac{1}{2\pi}$ proviene del componente sinusoidal y refleja el sesgo de la distribución
- Como la distribución le da más peso relativo a monedas con probabilidades menores (especialmente $p = 0.25$), el resultado global es menor que 0.5

Resolución 3.2

Imagina que inicialmente no sabemos nada específico sobre nuestra moneda, solo que viene de la máquina defectuosa. Pero cuando lanzamos la moneda y sale cara, esta observación nos da información valiosa

- **Antes del lanzamiento:** Cualquier moneda podría tener probabilidades desde $p = 0$ hasta $p = 1$, pero algunas son más probables según $f_P(p) = 1 + \sin(2\pi p)$
- **Después de observar cara:** Las monedas con p muy pequeño (digamos $p = 0.1$) se vuelven menos creíbles, mientras que las monedas con p más grande (digamos $p = 0.8$) se vuelven más creíbles
- Esto porque si una moneda tiene $p = 0.1$, es muy raro que dé cara, pero si tiene $p = 0.8$, es bastante común que dé cara

La situación cambia fundamentalmente cuando observamos que el primer lanzamiento dio cara. Esta nueva información nos permite actualizar nuestro conocimiento sobre la distribución de P usando inferencia bayesiana. La observación nos proporciona evidencia que modifica nuestra creencia inicial sobre qué valores de P son más probables.

Recordemos que la regla de Bayes viene dada por:

$$\mathbb{P}(H|E) = \frac{\mathbb{P}(E|H)\mathbb{P}(H)}{\mathbb{P}(E)} \tag{78}$$

donde H es una hipótesis (en este caso, un valor específico de P) y E es la evidencia observada (el evento de obtener cara). Usando la regla de Bayes para encontrar la distribución posterior de P dado que observamos cara:

$$f_{P|A}(p) = \frac{\mathbb{P}(A|P = p)f_P(p)}{\mathbb{P}(A)} \tag{79}$$

Donde:

- $\mathbb{P}(A|P = p) = p$ (si la moneda tiene probabilidad p, entonces la probabilidad de cara es p)
- $f_P(p) = 1 + \sin(2\pi p)$ (distribución original de la máquina)
- $\mathbb{P}(A) = \frac{\pi-1}{2\pi}$ (calculado en la parte anterior)

Sustituyendo:

$$f_{P|A}(p) = \frac{p \cdot (1 + \sin(2\pi p))}{\frac{\pi-1}{2\pi}} = \frac{2\pi p(1 + \sin(2\pi p))}{\pi - 1} \tag{80}$$

Por tanto:

$$f_{P|A}(p) = \begin{cases} \frac{2\pi p(1+\sin(2\pi p))}{\pi-1}, & \text{si } 0 \leq p \leq 1, \\ 0, & \text{en otro caso.} \end{cases} \tag{81}$$

Comparando $f_{P|A}(p)$ con la distribución original $f_P(p)$:

- La nueva distribución tiene un factor adicional de p en el numerador

- Esto significa que valores más grandes de p reciben mayor peso relativo
- Por ejemplo: si $p = 0.1$, el factor $p = 0.1$ reduce la probabilidad; si $p = 0.8$, el factor $p = 0.8$ la aumenta considerablemente
- Después de observar una cara, es más probable que nuestra moneda tenga una probabilidad alta de dar cara

Resolución 3.3

Ahora queremos predecir el segundo lanzamiento, pero con una ventaja: *ya sabemos que el primer lanzamiento dio cara.* Esta información cambia todo nuestro análisis.

- **Situación inicial:** Tenemos una moneda de la máquina defectuosa y el primer lanzamiento dio cara
- **Lo que sabemos ahora:** Gracias al primer lanzamiento, tenemos una mejor idea de qué tipo de moneda es (distribución actualizada $f_{P|A}(p)$)
- **Lo que queremos:** La probabilidad de que el segundo lanzamiento también sea cara
- **Clave importante:** Los lanzamientos son independientes *dado el valor de P*, pero como no conocemos P exactamente, hay dependencia estadística entre los lanzamientos

Para predecir el resultado del segundo lanzamiento, debemos utilizar toda la información disponible. Esto significa que ya no usamos la distribución original $f_P(p)$, sino la distribución actualizada $f_{P|A}(p)$ que incorpora la información del primer lanzamiento.

Es importante recordar que los lanzamientos son independientes condicionalmente a P, lo que significa que una vez que conocemos (o tenemos una distribución sobre) el valor de P, los resultados de diferentes lanzamientos no se influyen mutuamente. Tenemos:

$$\begin{aligned}
\mathbb{P}(B|A) &= \int_0^1 \mathbb{P}(B|P=p, A) f_{P|A}(p)\, dp && (82)\\
&= \int_0^1 \mathbb{P}(B|P=p) f_{P|A}(p)\, dp && (83)\\
&= \frac{2\pi}{\pi - 1} \int_0^1 p^2 (1 + \sin(2\pi p))\, dp. && (84)
\end{aligned}$$

Explicación de cada paso:

- **Primera línea:** Usamos el teorema de probabilidad total, pero ahora con la distribución actualizada $f_{P|A}(p)$
- **Segunda línea:** Aplicamos independencia condicional: $\mathbb{P}(B|P = p, A) = \mathbb{P}(B|P = p) = p$
- **¿Por qué independencia condicional?** Si sabemos que $P = p$, entonces el resultado del primer lanzamiento no afecta directamente al segundo; ambos solo dependen de p
- **Tercera línea:** Sustituimos $f_{P|A}(p) = \frac{2\pi p(1+\sin(2\pi p))}{\pi - 1}$

Observa que ahora tenemos p^2 en la integral (comparado con solo p en la parte (a)). Este p^2 proviene de:

- Un factor p de $\mathbb{P}(B|P = p) = p$ (probabilidad del segundo lanzamiento)
- Un factor p de $f_{P|A}(p)$ (la actualización bayesiana que favorece valores grandes de p)

Necesitamos calcular:

$$\int_0^1 p^2(1 + \sin(2\pi p))\, dp \tag{85}$$

Separamos la integral en dos partes:

$$\int_0^1 p^2(1 + \sin(2\pi p))\, dp = \int_0^1 p^2\, dp + \int_0^1 p^2 \sin(2\pi p)\, dp \tag{86}$$

Primera integral:

$$\int_0^1 p^2\, dp = \left[\frac{p^3}{3}\right]_0^1 = \frac{1^3}{3} - \frac{0^3}{3} = \frac{1}{3} \tag{87}$$

Segunda integral: Para $\int_0^1 p^2 \sin(2\pi p)\, dp$, usamos integración por partes dos veces.

$$u = p^2 \quad \Rightarrow \quad du = 2p\, dp \tag{88}$$

$$dv = \sin(2\pi p)\, dp \quad \Rightarrow \quad v = -\frac{1}{2\pi}\cos(2\pi p) \tag{89}$$

$$\int_0^1 p^2 \sin(2\pi p)\, dp = \left[p^2 \cdot \left(-\frac{1}{2\pi}\cos(2\pi p)\right)\right]_0^1 - \int_0^1 \left(-\frac{1}{2\pi}\cos(2\pi p)\right) \cdot 2p\, dp \tag{90}$$

$$= \left[-\frac{p^2}{2\pi}\cos(2\pi p)\right]_0^1 + \frac{1}{\pi}\int_0^1 p\cos(2\pi p)\, dp \tag{91}$$

Evaluando el primer término:

$$\left[-\frac{p^2}{2\pi}\cos(2\pi p)\right]_0^1 = -\frac{1^2}{2\pi}\cos(2\pi) - \left(-\frac{0^2}{2\pi}\cos(0)\right) \tag{92}$$

$$= -\frac{1}{2\pi} \cdot 1 - 0 = -\frac{1}{2\pi} \tag{93}$$

Segunda integración por partes para $\int_0^1 p\cos(2\pi p)\, dp$:

$$u = p \quad \Rightarrow \quad du = dp \tag{94}$$

$$dv = \cos(2\pi p)\, dp \quad \Rightarrow \quad v = \frac{1}{2\pi}\sin(2\pi p) \tag{95}$$

$$\int_0^1 p\cos(2\pi p)\,dp = \left[p \cdot \frac{1}{2\pi}\sin(2\pi p)\right]_0^1 - \int_0^1 \frac{1}{2\pi}\sin(2\pi p)\,dp \tag{96}$$

$$= \left[\frac{p}{2\pi}\sin(2\pi p)\right]_0^1 + \frac{1}{2\pi}\int_0^1 \sin(2\pi p)\,dp \tag{97}$$

$$= 0 + \frac{1}{2\pi}\left[-\frac{1}{2\pi}\cos(2\pi p)\right]_0^1 \tag{98}$$

$$= -\frac{1}{4\pi^2}[\cos(2\pi) - \cos(0)] = -\frac{1}{4\pi^2}[1-1] = 0 \tag{99}$$

Por tanto:

$$\int_0^1 p^2\sin(2\pi p)\,dp = -\frac{1}{2\pi} + \frac{1}{\pi}\cdot 0 = -\frac{1}{2\pi} \tag{100}$$

Combinando ambas integrales:

$$\int_0^1 p^2(1+\sin(2\pi p))\,dp = \int_0^1 p^2\,dp + \int_0^1 p^2\sin(2\pi p)\,dp \tag{101}$$

$$= \frac{1}{3} + \left(-\frac{1}{2\pi}\right) \tag{102}$$

$$= \frac{1}{3} - \frac{1}{2\pi} \tag{103}$$

$$= \frac{2\pi - 3}{6\pi} \tag{104}$$

Finalmente, sustituimos este resultado en la expresión para $\mathbb{P}(B|A)$:

$$\mathbb{P}(B|A) = \frac{2\pi}{\pi - 1}\cdot\frac{2\pi - 3}{6\pi} \tag{105}$$

$$= \frac{2\pi(2\pi - 3)}{6\pi(\pi - 1)} \tag{106}$$

$$= \frac{2(2\pi - 3)}{6(\pi - 1)} \tag{107}$$

$$= \frac{2\pi - 3}{3(\pi - 1)} \tag{108}$$

$$= \frac{2\pi - 3}{3\pi - 3} \approx 0.5110 \tag{109}$$

El resultado $0.5110 > 0.5$ confirma nuestra intuición y tiene mucho sentido:

- **Comparación:** Recuerda que $\mathbb{P}(A) = \frac{\pi-1}{2\pi} \approx 0.341 < 0.5$
- **Efecto del aprendizaje:** Después de observar cara, nuestra estimación sube de 0.341 a 0.511
- **¿Por qué es mayor que 0.5?** El primer lanzamiento nos dio evidencia de que probablemente tenemos una moneda con $p > 0.5$
- **Dependencia estadística:** Aunque los lanzamientos son físicamente independientes, estadísticamente están correlacionados a través de nuestro desconocimiento de P

4. El negocio "Donde la Sonia" tiene N clientes regulares, donde N es una variable aleatoria con PMF

$$p_N(n) = p^{n-1}(1-p) \quad \text{para } n = 1, 2, 3, \ldots \tag{110}$$

Cada viernes por la noche, el negocio organiza una promoción especial. Cada cliente regular decide asistir a la promoción con probabilidad q, independientemente de todos los otros clientes. Si un cliente asiste a la promoción, entonces gasta una cantidad de dinero M, que es una variable aleatoria continua con PDF

$$f_M(m) = \lambda e^{-\lambda m} \quad \text{para } m \geq 0. \tag{111}$$

N, M, y si cada cliente asiste son todos independientes. Determine:

1. La esperanza y varianza del número de clientes que asisten a la promoción.
2. [**Propuesto**]La esperanza y varianza de la cantidad total de dinero gastada en la promoción.

Solución:

Resolución 4.1

Este problema presenta una situación de aleatoriedad compuesta que es muy común en la vida real. Imagina la situación del negocio "Donde la Sonia": tenemos dos fuentes de incertidumbre que se combinan. Primero, no sabemos cuántos clientes regulares tiene el negocio, este número N es aleatorio y sigue una distribución geométrica. Segundo, cada cliente decide independientemente si asistir o no a la promoción con probabilidad q. El resultado final es que el número de asistentes K depende de ambas fuentes de incertidumbre, creando una estructura de dependencia compleja.

La distribución geométrica $p_N(n) = p^{n-1}(1-p)$ modela situaciones donde contamos "hasta el primer éxito". En nuestro contexto, podemos interpretarla como el número de clientes que el negocio puede mantener antes de que algún factor externo (competencia, cambio de ubicación, etc.) cause una "pérdida".

Primero identifiquemos todas nuestras variables aleatorias y sus propiedades:

- N: Número de clientes regulares (distribución geométrica)
- B: Variable Bernoulli que indica si un cliente asiste ($B = 1$) o no ($B = 0$)
- M: Cantidad gastada por un cliente que asiste (distribución exponencial)
- K: Número total de clientes que asisten a la promoción

Las propiedades de estas distribuciones son:

$$\mathbb{E}[N] = \frac{1}{1-p}, \qquad \text{var}(N) = \frac{p}{(1-p)^2}, \tag{112}$$

$$\mathbb{E}[M] = \frac{1}{\lambda}, \qquad \text{var}(M) = \frac{1}{\lambda^2}, \tag{113}$$

$$\mathbb{E}[B] = q, \qquad \text{var}(B) = q(1-q). \tag{114}$$

Para entender mejor estos valores, consideremos algunos ejemplos: si $p = 0.1$, entonces $\mathbb{E}[N] = \frac{1}{0.9} \approx 1.11$ clientes en promedio; si $p = 0.5$, entonces $\mathbb{E}[N] = 2$ clientes en promedio; y si $p = 0.9$, entonces $\mathbb{E}[N] = 10$ clientes en promedio. Notemos que la varianza crece considerablemente cuando p se acerca a 1.

El número de clientes que asisten es una suma aleatoria: $K = B_1 + B_2 + \cdots + B_N$. Esto significa que cada B_i es una variable Bernoulli independiente (el cliente i asiste con probabilidad q), pero el número de términos N también es aleatorio. Esta es la esencia de una "suma aleatoria": sumamos un número aleatorio de variables aleatorias.

Una suma aleatoria tiene la forma general $S = X_1 + X_2 + \cdots + X_N$, donde N es una variable aleatoria independiente de las X_i, y todas las X_i son independientes e idénticamente distribuidas. Para entender de dónde salen las fórmulas, consideremos que condicionalmente a $N = n$, tenemos una suma fija $S|N = n = X_1 + \cdots + X_n$.

Usando la ley de expectativas totales, la esperanza de una suma aleatoria es:

$$\mathbb{E}[S] = \mathbb{E}[\mathbb{E}[S|N]] = \mathbb{E}[\mathbb{E}[X_1 + \cdots + X_N|N]] \tag{115}$$
$$= \mathbb{E}[N \cdot \mathbb{E}[X_1]] = \mathbb{E}[N] \cdot \mathbb{E}[X_1] \tag{116}$$

donde usamos que condicionalmente a $N = n$, la suma es determinística y $\mathbb{E}[X_1 + \cdots + X_n] = n \cdot \mathbb{E}[X_1]$. Esta es la fórmula general para la esperanza de sumas aleatorias.

En nuestro caso particular, aplicando esta fórmula:

$$\mathbb{E}[K] = \mathbb{E}[N] \cdot \mathbb{E}[B] = \frac{1}{1-p} \cdot q = \frac{q}{1-p} \tag{117}$$

La interpretación es simple: si tenemos en promedio $\frac{1}{1-p}$ clientes y cada uno asiste con probabilidad q, entonces esperamos $\frac{q}{1-p}$ asistentes.

Para la varianza de sumas aleatorias, la derivación es más compleja. Usando la ley de varianzas totales $\text{var}(S) = \mathbb{E}[\text{var}(S|N)] + \text{var}(\mathbb{E}[S|N])$:

$$\text{var}(S) = \mathbb{E}[\text{var}(X_1 + \cdots + X_N|N)] + \text{var}(N \cdot \mathbb{E}[X_1]) \tag{118}$$
$$= \mathbb{E}[N \cdot \text{var}(X_1)] + (\mathbb{E}[X_1])^2 \cdot \text{var}(N) \tag{119}$$
$$= \mathbb{E}[N] \cdot \text{var}(X_1) + (\mathbb{E}[X_1])^2 \cdot \text{var}(N) \tag{120}$$

Esta es la fórmula general para la varianza de sumas aleatorias. En nuestro problema específico:

$$\text{var}(K) = \mathbb{E}[N] \cdot \text{var}(B) + (\mathbb{E}[B])^2 \cdot \text{var}(N) \tag{121}$$

Sustituyendo valores:

$$\text{var}(K) = \frac{1}{1-p} \cdot q(1-q) + q^2 \cdot \frac{p}{(1-p)^2} \tag{122}$$
$$= \frac{q(1-q)}{1-p} + \frac{pq^2}{(1-p)^2} \tag{123}$$

La fórmula de varianza para sumas aleatorias tiene dos componentes que reflejan las dos fuentes de incertidumbre:

Primer término: $\mathbb{E}[N] \cdot \mathbf{var}(B) = \frac{q(1-q)}{1-p}$ - Este representa la variabilidad inherente en las decisiones individuales. Incluso si supiéramos exactamente cuántos clientes hay (digamos $N = n$ fijo), cada cliente decide independientemente si asistir, creando variabilidad. Si tenemos n clientes, la varianza sería $n \cdot q(1-q)$. Como no sabemos n exactamente, tomamos su valor esperado.

Segundo término: $(\mathbb{E}[B])^2 \cdot \mathbf{var}(N) = \frac{pq^2}{(1-p)^2}$ - Este representa la variabilidad adicional debido a no saber cuántos clientes tenemos. Si cada cliente asiste con probabilidad q en promedio, pero el número total de clientes N es incierto, esto amplifica la variabilidad total. El factor q^2 aparece porque la incertidumbre en N se propaga cuadráticamente a través de la esperanza condicional.

Como ejemplo numérico, si $p = 0.5$ y $q = 0.7$, entonces $\mathbb{E}[K] = \frac{0.7}{1-0.5} = 1.4$ clientes esperados y $\text{var}(K) = \frac{0.7 \cdot 0.3}{0.5} + \frac{0.5 \cdot 0.7^2}{0.5^2} = 0.42 + 0.98 = 1.4$.

Resolución 4.2

El problema se vuelve aún más fascinante cuando consideramos el dinero total gastado en la promoción. Ahora tenemos una cadena de dependencias más compleja: $N \to K \to G$, donde cada eslabón introduce su propia variabilidad. Primero, el número aleatorio de clientes regulares N determina quiénes pueden asistir. Segundo, cada cliente decide independientemente si asistir, dando lugar al número aleatorio de asistentes K. Finalmente, cada asistente gasta una cantidad aleatoria de dinero M, resultando en el gasto total G.

Esta situación es común en muchos contextos reales. Por ejemplo, en "Donde la Sonia", no solo es incierto cuántos clientes asistirán a la promoción, sino que también es incierto cuánto gastará cada uno. Algunos clientes pueden venir solo a mirar y gastar poco, mientras que otros pueden aprovechar las ofertas y gastar considerablemente.

Sea G el dinero total gastado en la promoción. Entonces $G = M_1 + M_2 + \cdots + M_K$, donde cada M_i es el gasto del i-ésimo cliente que asiste. Notemos que esto es nuevamente una suma aleatoria, pero ahora tenemos una suma aleatoria anidada: el número de términos K es en sí mismo el resultado de otra suma aleatoria.

La distribución exponencial $f_M(m) = \lambda e^{-\lambda m}$ para el gasto individual es muy realista en contextos comerciales. Esta distribución tiene la propiedad de que la mayoría de los clientes gastan cantidades pequeñas (cerca de 0), pero ocasionalmente algunos clientes gastan cantidades muy grandes. El parámetro λ controla qué tan "generosos" son los clientes en promedio: valores grandes de λ significan que los gastos tienden a ser pequeños, mientras que valores pequeños de λ permiten gastos más grandes.

Para calcular la esperanza del gasto total, aplicamos nuevamente la fórmula para sumas aleatorias:

$$\mathbb{E}[G] = \mathbb{E}[M] \cdot \mathbb{E}[K] = \frac{1}{\lambda} \cdot \frac{q}{1-p} = \frac{q}{\lambda(1-p)} \tag{124}$$

Esta fórmula tiene una interpretación muy clara: el gasto promedio por cliente ($\frac{1}{\lambda}$) multiplicado por el número promedio de clientes que asisten ($\frac{q}{1-p}$) nos da el gasto total promedio.

Para la varianza, necesitamos la fórmula para sumas aleatorias, que considera tanto la variabilidad en los gastos individuales como la variabilidad en el número de gastadores:

$$\operatorname{var}(G) = \operatorname{var}(M) \cdot \mathbb{E}[K] + (\mathbb{E}[M])^2 \cdot \operatorname{var}(K) \tag{125}$$

$$= \frac{1}{\lambda^2} \cdot \frac{q}{1-p} + \left(\frac{1}{\lambda}\right)^2 \cdot \left(\frac{q(1-q)}{1-p} + \frac{pq^2}{(1-p)^2}\right) \tag{126}$$

$$= \frac{q}{\lambda^2(1-p)} + \frac{1}{\lambda^2}\left(\frac{q(1-q)}{1-p} + \frac{pq^2}{(1-p)^2}\right) \tag{127}$$

La varianza total tiene una estructura muy instructiva. El primer término $\frac{q}{\lambda^2(1-p)}$ representa la variabilidad inherente en los gastos individuales: incluso si supiéramos exactamente cuántos clientes van a asistir, habría incertidumbre sobre cuánto gastará cada uno. El segundo término $\frac{1}{\lambda^2}\left(\frac{q(1-q)}{1-p} + \frac{pq^2}{(1-p)^2}\right)$ representa la variabilidad adicional debido a no saber cuántos clientes asistirán. Este segundo término toma toda la variabilidad que calculamos para K en la parte anterior y la amplifica por el cuadrado del gasto promedio.

Como ejemplo ilustrativo, supongamos que $p = 0.6$, $q = 0.8$ y $\lambda = 0.1$. Entonces el gasto promedio esperado sería $\mathbb{E}[G] = \frac{0.8}{0.1(1-0.6)} = \frac{0.8}{0.04} = 20$ unidades monetarias. La varianza sería considerable debido a la combinación de todas las fuentes de incertidumbre, mostrando que el negocio debe estar preparado para una amplia gama de ingresos posibles durante la promoción.