T1.1 Razonamiento probabilístico: representación e inferencia

Índice

1 El problema de la calificación

Problema de la calificación: imposibilidad práctica de conocer y comprobar todas las **calificaciones** (condiciones) que habría que garantizar para asegurar el cumplimiento de una acción

- Ejemplo: salir al aeropuerto 90 minutos antes del vuelo me permite llegar a tiempo SI no hay atascos Y no hay pinchazos Y ...
- Ejemplo: un bote nos permite cruzar un río SI es un bote de remo Y tiene remos y escálamos Y no están rotos Y encajan Y ...

Incertidumbre: los sistemas inteligentes actuales incluyen la **incertidumbre** como parte del conocimiento y la representan mediante **probabilidades** asociadas a los sucesos (proposiciones) de interés

2 Representación probabilística

Distribución de probabilidad conjunta: de las variables aleatorias de interés para representar el conocimiento probabilístico

Ejemplo del dentista: conocimiento para diagnosticar caries

Variables aleatorias de interés:

$$\begin{array}{ll} \text{Dolor:} & D \in \{0,1\} \\ \text{Caries:} & C \in \{0,1\} \\ \text{Hueco:} & H \in \{0,1\} \end{array}$$

Representación: tabla a la derecha con

$$P(D=d, C=c, H=h) \quad \text{para todo } d, c, h \in \{0,1\}$$

$1000 \to 576$

	d	c	$\cancel{h}$ h	P
	0	0	0	0.576
	0	0	1	0.008
	0	1	0	0.144
$d=$	0	1	1	0.072
	1	0	0	0.064
	1	0	1	0.012
	1	1	0	0.016
$d=$	1	1	1	0.108

3 Inferencia probabilística

$P(y, x, z)$

Reglas suma y producto: reglas básicas para calcular la probabilidad de cualquier **suceso** (**proposición**) de interés a partir de la distribución conjunta

$P(x,y,z) = P(x \mid y,z) \cdot P(y,z) = P(x \mid y,z) \cdot P(y \mid z) \cdot P(z)$

$= P(x \mid y,z) \cdot P(z,y) \cdot P(y)$

$$\boxed{P(x) = \sum_y P(x, y)} \qquad \text{y} \qquad P(x, y) = P(x)\, P(y \mid x)$$

$= P(y) \cdot P(x \mid y)$

Observación importante: en general no es necesario conocer la tabla completa de probabilidades conjuntas para calcular la probabilidad de un suceso dado mediante las reglas suma y producto

$= 0.072 + 0.108$

Ejemplo del dentista: cálculo de la probabilidad de observar...

$= P(d=0, c=1, h=1) + P(d=1, c=1, h=1)$

- Caries y hueco (a la vez): $P(c = 1, h = 1) = \sum_{d=0,1} P(d, c = 1, h = 1) = 0.180$
- Hueco: $P(h = 1) = \sum_{d=0,1} \sum_{c=0,1} P(d, c, h = 1) = 0.200$ $= P(d=0, c=0, h=1) + P(d=0, c=1, h=1) + P(d=1, c=0, h=1) + P(d=1, c=1, h=1)$
- Caries después de observar hueco: $P(c = 1 \mid h = 1) = \boxed{\dfrac{P(c = 1, h = 1)}{P(h = 1)}} = \dfrac{0.180}{0.200} = 0.900$

In [1]:
```
import numpy as np
T = np.array([[0,0,0,.576], [0,0,1,.008], [0,1,0,.144], [0,1,1,.072],
              [1,0,0,.064], [1,0,1,.012], [1,1,0,.016], [1,1,1,.108]])
Pc1b1 = np.sum(T[(T[:,1]==1) & (T[:,2]==1),-1])
Pb1 = np.sum(T[T[:,2]==1,-1])
Pc1Db1 = Pc1b1/Pb1
print(f"Pc1b1 = {Pc1b1:.3f}  Pb1 = {Pb1:.3f}  Pc1Db1 = {Pc1Db1:.3f}")
```

```
Pc1b1 = 0.180  Pb1 = 0.200  Pc1Db1 = 0.900
```

$P(c=1, h=1) = \boxed{P(c=1 \mid h=1)}\, P(h=1)$

4 Independencia

Variables independientes: dos variables x y y son **independientes** si

$$P(x,y) = P(x)\,P(y) \quad \text{o} \quad P(x \mid y) = P(x) \quad \text{o} \quad P(y \mid x) = P(y)$$

Conocimiento experto: la independencia puede establecerse por conocimiento experto y conveniencia

Ejemplo del dentista:

- Consideramos una nueva variable con el tiempo que hace cuando el paciente visita el dentista

$$T \in \{\text{sol}, \text{nubes}, \text{lluvia}, \text{nieve}\}$$

- Asumimos que las tres variables que ya teníamos son independientes del tiempo que hace

$$\boxed{P(d,c,h,t)} = P(t)\,P(d,c,h \mid \cancel{t}) = P(t)\,P(d,c,h)$$

- Así reducimos el número de probabilidades a almacenar: 32 vs $4+8$

5 Teorema de Bayes

Teorema de Bayes: permite actualizar nuestro conocimiento sobre una hipótesis y después de observar una nueva evidencia x

$$P(x,y) = \boxed{P(y \mid x)} \cdot P(x)$$

$$\boxed{P(y \mid x) = \frac{P(x,y)}{P(x)}} = \boxed{P(y)} \frac{P(x \mid y)}{P(x)} = \frac{P(x \mid y) \cdot P(y)}{P(x)}$$

- De otra manera: $P(y \mid x)$ es la probabilidad de que se produzca el efecto y después de observar que se ha producido la causa x

Ejemplo del dentista:

- Sabemos que la probabilidad de caries es: $P(c=1) = 0.34$
- Sabemos que la probabilidad de dolor es: $P(d=1) = 0.20$
- Sabemos que la probabilidad de dolor después de observar caries es: $P(d=1 \mid c=1) = 0.36$
- ¿Cúal es la probabilidad de caries después de observar dolor, $\boxed{P(c=1 \mid d=1)}$?

$$P(c=1 \mid d=1) = P(c=1)\,\frac{P(d=1 \mid c=1)}{P(d=1)} = 0.34\,\frac{0.36}{0.20} = 0.61$$

In [3]:
```python
Pc1 = 0.34; Pd1 = 0.20; Pd1c1 = 0.36; Pc1Dd1 = Pc1 * Pd1c1 / Pb1; print(f"Pc1Dd1 = {Pc1Dd1:.2f}")
```

```
Pc1Dd1 = 0.61
```

$$P(c=1 \mid d=1) = \frac{P(c=1, d=1)}{P(d=1)} = \frac{P(d=1 \mid c=1) \cdot P(c=1)}{P(d=1)}$$

$$P(c=1, d=1) = P(c=1 \mid d=1)\, P(d=1)$$

Ejercicios T1.1 Razonamiento probabilístico

2023_01_26_Cuestión 1: Dada la siguiente tabla de probabilidades conjuntas de tres variables de interés:

A	0	0	0	0	1	1	1	1
B	0	0	1	1	0	0	1	1
C	0	1	0	1	0	1	0	1
$P(A,B,C)$	0.093	0.100	0.133	0.163	0.157	0.150	0.117	0.087

Cúal és el valor de $P(A=1, B=1 \mid C=1)$?

1. $P(A=1, B=1 \mid C=1) \leq 0.25$
2. $0.25 < P(A=1, B=1 \mid C=1) \leq 0.50$
3. $0.50 < P(A=1, B=1 \mid C=1) \leq 0.75$
4. $0.75 < P(A=1, B=1 \mid C=1)$

$$P(A=1, B=1 \mid C=1) = \frac{P(A=1, B=1, C=1)}{P(C=1)} = \frac{0.087}{0.05} = 0.174$$

$$P(C=1) = \sum_{A\in\{0,1\}} \sum_{B\in\{0,1\}} P(A,B,C=1) = P(A=0,B=0,C=1)+P(A=0,B=1,C=1)+P(A=1,B=0,C=1)+P(A=1,B=1,C=1)$$
$$= 0.5$$

$$P(A=1,B=1,C=1) = \boxed{P(A=1,B=1 \mid C=1)} \cdot P(C=1)$$

Solución: la 1; $P(A=1, B=1 \mid C=1) = 0.174$

In [1]:
```
PA1B1C1 = 0.087; PC1 = 0.1 + 0.163 + 0.15 + 0.087; PA1B1_C1 = PA1B1C1 / PC1; print(PA1B1_C1)
```

```
0.174
```

2023_01_17_Cuestión 1: Supón que tenemos dos cajas con 40 naranjas en la primera y 80 en la segunda. La primera caja contiene 9 naranjas Navelina y 31 Caracara. La segunda caja contiene tres veces más naranjas Navelina que Caracara. Ahora supón que se escoge una caja al azar, y después una naranja al azar de la caja escogida. Si la naranja escogida es Navelina, la probabilidad P de que proceda de la primera caja es:

1. $0/4 \leq P < 1/4$.
2. $1/4 \leq P < 2/4$.
3. $2/4 \leq P < 3/4$.
4. $3/4 \leq P \leq 4/4$.

$C = \{1, 2\}$

$O = \{N, C\}$

$P(C=1) = \frac{40}{40+80} = \frac{1}{3} = 0.33$

$P(C=2) = \frac{80}{40+80} = \frac{2}{3}$

$P(O=N \mid C=1) = \frac{9}{40}$ $\quad P(O=N \mid C=2) = \frac{3}{4}$

$P(O=C \mid C=1) = \frac{31}{40}$ $\quad P(O=C \mid C=2) = \frac{1}{4}$

$3 \cdot x + x = 1$

$$P(C=1 \mid O=N) = \frac{P(O=N \mid C=1) \cdot P(C=1)}{P(O=N)}$$

$$P(O=N) = \sum_{C \in \{1,2\}} P(C, O=N) = \sum_{C \in \{1,2\}} P(O=N \mid C) \cdot P(C) = P(O=N \mid C=1) \cdot P(C=1) + P(O=N \mid C=2) \cdot P(C=2)$$

Solución:

$$
\begin{aligned}
P = P(C=1 \mid T=N) &= \frac{P(C=1)P(T=N \mid C=1)}{P(C=1)P(T=N \mid C=1) + P(C=2)P(T=N \mid C=2)} \\
&= \frac{1/2 \cdot 9/40}{1/2 \cdot 9/40 + 1/2 \cdot 3/4} = \frac{9}{9+30} = 0.23
\end{aligned}
$$

2022_01_27_Cuestión 4: Dada la siguiente tabla de frecuencias conjuntas de tres variables de interés:

A	0	0	0	0	1	1	1	1
B	0	0	1	1	0	0	1	1
C	0	1	0	1	0	1	0	1
$N(A, B, C)$	124	28	227	175	126	222	23	75

Cúal és el valor de $P(A = 1 \mid B = 1, C = 0)$?

1. 0.023
2. 0.250
3. 0.092
4. 0.446

Solución: la 3.

2022_01_13_Cuestión 7: Sea un problema de razonamiento probabilístico sobre desplazamientos por carretera, con las variables aleatorias de interés:

- Climatología (C): {claro(CLA), nublado (NUB), lluvia (LLU)}
- Luminosidad (L): {día (DÍA), noche (NOCHE)}
- Seguridad (S): {seguro (SEG), accidente (ACC)}

La probabilidad conjunta de las tres variables viene dada en la tabla:

	DÍA			NOCHE		
$P(S, L, C)$	CLA	NUB	LLU	CLA	NUB	LLU
SEG	0.27	0.23	0.07	0.16	0.07	0.06
ACC	0.02	0.01	0.02	0.02	0.03	0.04

La probabilidad condicional $P(S = \mathrm{ACC} \mid L = \mathrm{DÍA}, C = \mathrm{NUB})$ es:

1. 0.042
2. 0.010
3. 0.240
4. 0.140

Solución: la 1.

T1.2 Variables continuas y regla de Bayes

Índice

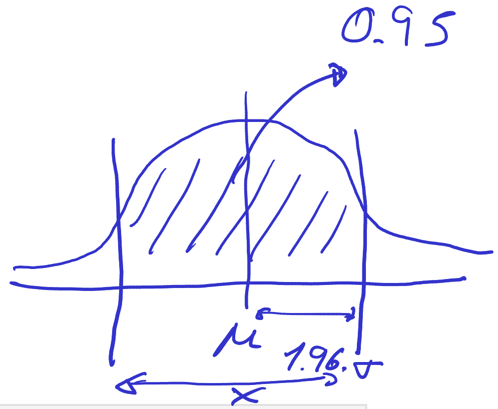
1 Variables continuas

Función de densidad de probabilidad: caracterización usual de las variables continuas para la representación de conocimiento probabilístico

$$p(x) \geq 0 \quad \text{para todo } x \qquad \text{y} \qquad \int p(x)\, dx = 1$$

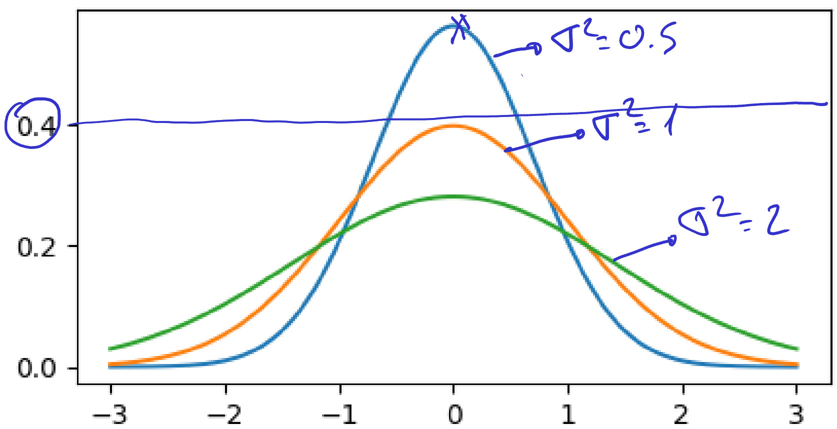
La densidad normal: $p(x) \sim \mathcal{N}(\mu, \sigma^2)$

$$\boxed{p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\,\sigma^2}\right)} \qquad P(x \in [\mu \pm 1.96\sigma]) = 0.95$$



Ejemplo: densidades normales con $\mu = 0$ y $\sigma^2 = 0.5, 1, 2$

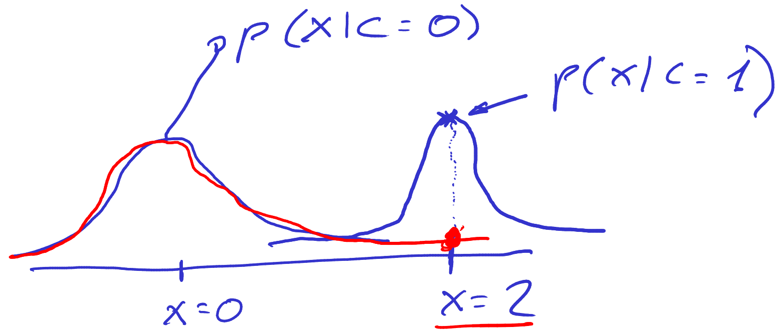
In [2]:
```python
import numpy as np; from scipy.stats import norm; import matplotlib.pyplot as plt
x = np.linspace(-3, 3, 200)
plt.figure(figsize=(5, 2.5))
plt.plot(x, norm.pdf(x, 0, np.sqrt(0.5)), x, norm.pdf(x, 0, 1), x, norm.pdf(x, 0, np.sqrt(2)));
```



2 Teorema de Bayes en el caso continuo

Teorema de Bayes en el caso continuo: probabilidad de una hipótesis y después de observar una evidencia (nueva) x

$$P(y \mid x) = P(y)\,\frac{p(x \mid y)}{p(x)}$$



Ejemplo: $x =$ resultado de un test de saliva para el diagnóstico de caries

- Sin caries, $c = 0$, $p(x \mid c = 0) \sim \mathcal{N}(\mu = 0, \sigma^2 = 1)$
- Con caries, $c = \cancel{0}\,1$, $p(x \mid c = 1) \sim \mathcal{N}(\mu = 2, \sigma^2 = 0.5)$
- Sabemos que la probabilidad (a priori) de caries es: $P(c = 1) = 0.34$
- **Si el test da $x = 2$, cuál es la probabilidad (a posteriori) de caries?**

$\boxed{P(c=0)} + P(c=1) = 1.0$

$$P(c = 1 \mid x = 2) = P(c = 1)\,\frac{p(x = 2 \mid c = 1)}{p(x = 2)} = 0.340\,\frac{0.564}{0.307} = 0.843$$

- Observa que primero había que encontrar la (densidad de) probabilidad (a priori) de test $x = 2$:

$$\begin{aligned} p(x = 2) &= P(c = 0)\,p(x = 2 \mid c = 0) + P(c = 1)\,p(x = 2 \mid c = 1) \\ &= (1 - 0.34) \cdot 0.054 + 0.34 \cdot 0.564 = 0.227 \end{aligned}$$

In [3]:
```python
Pc1 = 0.34; px2Dc0 = norm.pdf(2, 0, 1); px2Dc1 = norm.pdf(2, 2, np.sqrt(0.5))
px2 = (1-Pc1) * px2Dc0 + Pc1 * px2Dc1; Pc1Dx2 = Pc1 * px2Dc1 / px2
print(f"px2Dc0 = {px2Dc0:.3f}  px2Dc1 = {px2Dc1:.3f}  px2 = {px2:.3f}  Pc1Dx2 = {Pc1Dx2:.3f}")
```

```
px2Dc0 = 0.054  px2Dc1 = 0.564  px2 = 0.227  Pc1Dx2 = 0.843
```

3 La regla de decisión de Bayes

clase imagen

Regla de decisión de Bayes: predice una hipótesis después de observar una evidencia x mediante la elección, entre un conjunto de hipótesis posibles $\mathcal{C}$, de una hipótesis de máxima **probabilidad a posteriori** (de la observación de la evidencia)

6 ó 9

$$c^*(x) = \operatorname*{argmax}_{c \in \mathcal{C}} \boxed{P(c \mid x)}$$

$P(c=6 \mid x = \boxed{6}) = 0.9$ $\quad c^*(x) = 6$

$P(c=9 \mid x \boxed{6}) = 0.1$

Probabilidad de error: es decir, probabilidad de que la hipótesis predicha sea distinta de la realmente producida

$$P(\text{error} \mid x) = 1 - P(c^*(x) \mid x)$$

$c^*(x) = 6 \to P(\text{error} \mid x) = 1 - 0.9 = 0.1$

$c^*(x) = 9 \to P(\text{error} \mid x) = 1 - 0.1 = 0.9$

Optimalidad de la regla de Bayes: ninguna otra elección mejoraría esta probabilidad de error!

Ejemplo del dentista:

$$c^*(x=2) = \boxed{\operatorname*{argmax}_{c}} \begin{pmatrix} P(c=0 \mid x=2) = 0.116 \\ P(c=1 \mid x=2) = 0.884 \end{pmatrix} = 1$$

Regla de Bayes en función de probabilidades a priori y (densidades) condicionales de las clases: en lugar (de arg-)maximizar $P(c \mid x)$ en c, lo hacemos en función de $P(c)\,p(x \mid c)$ puesto que el resultado es el mismo

$$\boxed{c^*(x) = \operatorname*{argmax}_{c \in \mathcal{C}} P(c \mid x)} = \operatorname*{argmax}_{c \in \mathcal{C}} \boxed{P(c)\,\frac{p(x \mid c)}{p(x)}} = \operatorname*{argmax}_{c \in \mathcal{C}} \boxed{P(c)\,p(x \mid c)}$$

Ejemplo del dentista:

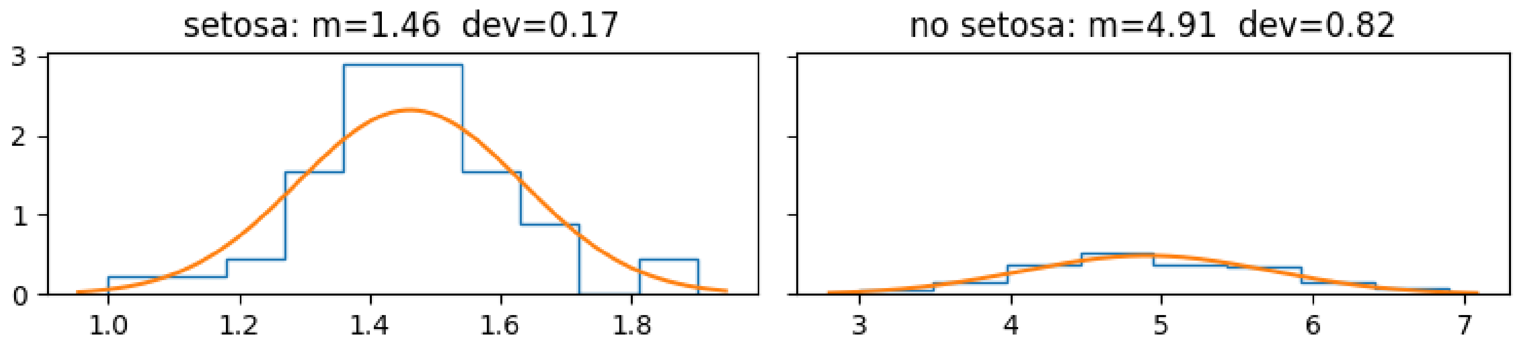
$$c^*(x=2) = \operatorname*{argmax}_{c} \begin{pmatrix} P(c=0)\,p(x=2 \mid c=0) = 0.036 \\ P(c=1)\,p(x=2 \mid c=1) = 0.271 \end{pmatrix} = 1$$

Ejercicios T1.2 Variables continuas y regla de Bayes

Problema: Considerad la clasificación de flores iris en setosa o no-setosa a partir de la longitud de pétalos, x. El estudio empírico siguiente muestra que las distribuciones de x para setosas y no-setosas pueden aproximarse con distribuciones normales de medias y desviaciones estándares:

$$p(x \mid c = \text{set}) \sim \mathcal{N}(\mu_{\text{set}} = 1.46, \sigma_{\text{set}} = 0.17) \qquad \text{y} \qquad p(x \mid c = \text{nos}) \sim \mathcal{N}(\mu_{\text{nos}} = 4.91, \sigma_{\text{nos}} = 0.82)$$

In [1]:
```python
import numpy as np; import matplotlib.pyplot as plt
from sklearn.datasets import load_iris; from scipy.stats import norm
iris = load_iris(); X = iris.data.astype(np.float16); y = iris.target.astype(np.uint)
x_set = np.squeeze(X[np.where(y==0), 2]); x_nos = np.squeeze(X[np.where(y!=0), 2])
fig, axs = plt.subplots(1, 2, figsize=(8, 2), sharey=True, tight_layout=True)
axs[0].hist(x_set, bins='auto', density=True, histtype='step')
x_set_range = np.arange(*axs[0].get_xlim(), .01)
x_set_mean = x_set.mean(); x_set_dev = np.sqrt(x_set.var())
axs[0].set_title(f'setosa: m={x_set_mean:.2f}  dev={x_set_dev:.2f}')
axs[0].plot(x_set_range, norm.pdf(x_set_range, x_set_mean, x_set_dev))
axs[1].hist(x_nos, bins='auto', density=True, histtype='step')
x_nos_range = np.arange(*axs[1].get_xlim(), .01)
x_nos_mean = x_nos.mean(); x_nos_dev = np.sqrt(x_nos.var())
axs[1].set_title(f'no setosa: m={x_nos_mean:.2f}  dev={x_nos_dev:.2f}')
axs[1].plot(x_nos_range, norm.pdf(x_nos_range, x_nos_mean, x_nos_dev));
```



Si las densidades normales estimadas son ciertas y la probabilidad a priori de setosa es $1/3$, ¿cuál es la probabilidad a posteriori de que una flor de longitud de pétalos 2 sea setosa?

Solución:

$$
\begin{aligned}
P(c=\text{set}\mid x=2) &= \frac{P(c=\text{set})\,p(x=2\mid c=\text{set})}{p(x=2)}\\
&= \frac{P(c=\text{set})\,p(x=2\mid c=\text{set})}{P(c=\text{set})\,p(x=2\mid c=\text{set})+P(c=\text{nos})\,p(x=2\mid c=\text{nos})}\\
&= \frac{1/3\cdot\mathcal{N}(x=2\mid\mu_{\text{set}}=1.46,\sigma_{\text{set}}=0.17)}{1/3\cdot\mathcal{N}(x=2\mid\mu_{\text{set}}=1.46,\sigma_{\text{set}}=0.17)+2/3\cdot\mathcal{N}(x=2\mid\mu_{\text{nos}}=4.91,\sigma_{\text{nos}}=0.82)}\\
&= \frac{\frac{1}{0.17}\exp\left(-\frac{(2-1.46)^2}{2\cdot 0.17^2}\right)}{\frac{1}{0.17}\exp\left(-\frac{(2-1.46)^2}{2\cdot 0.17^2}\right)+\frac{2}{0.82}\exp\left(-\frac{(2-4.91)^2}{2\cdot 0.82^2}\right)} = \frac{0.0379}{0.0379+0.0045} = 0.89
\end{aligned}
$$