ความน่าจะเป็นและสัญญาณสุ่ม (Probability & Random Signal)

ตอนที่ 1 : ความไม่แน่นอนในกรณีที่แย่ที่สุด (Worst case Uncertainty)

ความไม่แน่นอน (uncertainty) เห็นพบได้ในระบบทางกายภาพทั่วไป บางระบบนั้นอาจจะได้รับผลกระทบจากความไม่แน่นอนมากกว่าระบบอื่น นั่นหมายความว่าผลตอบสนองของระบบนั้นสามารถเป็นได้หลากหลาย เราจึงจำเป็นต้องระบุความไม่แน่นอนในระบบเท่าที่ทำได้ และ ออกแบบตัวกรอง ระบบประมาณค่า หรือ ระบบควบคุม ที่จัดการความไม่แน่นอนนั้น

หนึ่งค่าที่บ่งบอกถึงลักษณะสำคัญของระบบคือพารามิเตอร์ (parameter) ซึ่งเป็นค่าคงที่ในระบบ ปัญหาในการวิเคราะห์หรือการออกแบบคือการที่เราไม่ทราบค่าที่แน่ชัดของค่าคงที่เหล่านี้ ดังนั้นเราสามารถเขียนค่าคงที่และบอกขอบเขตความไม่แน่นอนของค่าเหล่าได้ดังนี้

$$e = \hat{e} \pm \delta e$$

โดยที่

e คือค่าของพารามิเตอร์

$\hat{e}$ คือค่าประมาณที่ดีที่สุด (best estimate)

δe คือค่าความไม่แน่นอนสัมบูรณ์ที่มีค่ามากที่สุดซึ่งมีหน่วยเดียวกันกับค่า $\hat{e}$ ซึ่งจะเป็นค่าบวกเสมอ

ความไม่แน่นอนที่ดเขียนในรูปบวกลบนั้นระบุถึงกรณีที่แย่ที่สุดที่ (worst case) จะเกิดขึ้นได้จริง ดังนั้นค่าที่มีความไม่แน่นอนนี้จะอยู่ในช่วง $[\hat{e} - \delta\, e, \hat{e} + \delta\, e]$

การอธิบายความไม่แน่นอนในลักษณะดังกล่าวมีข้อดีในเรื่องของการออกแบบและวิเคราะห์ให้ครอบคลุม เราสามารถออกแบบระบบที่คงทน (robust) และทำงานอยู่ในช่วงได้

การบวกลบกันของความไม่แน่นอน

กำหนดให้ชิ้นงาน 2 ชิ้นวางต่อกันโดยที่ชิ้นงานแรกมีความยาวเท่ากับ $\hat{e}_1 \pm \delta e_1$ ชิ้นงานอีกชิ้นหนึ่งมีความยาวเท่ากับ $\hat{e}_2 \pm \delta e_2$ ดังนั้นหากเราเอาชิ้นงานวางต่อกัน ความยาวสูงสุดของชิ้นงานที่เป็นไปได้ $e_{\max}$ สามารถเขียนได้ดังต่อไปนี้

$$e_{\max} = \left(\hat{e}_1 + \delta e_1\right) + \left(\hat{e}_2 + \delta e_2\right)$$

ความยาวต่ำสุดของชิ้นงานที่เป็นไปได้ $e_{\min}$ เป็นดังต่อไปนี้

$$e_{\min} = \left(\hat{e}_1 - \delta e_1\right) + \left(\hat{e}_2 - \delta e_2\right)$$

หากเราเขียนอยู่ในรูป $\pm$ เราจะสามารถเขียนได้ดังต่อไปนี้

$$\begin{aligned}
e &= \left(\hat{e}_1 + \hat{e}_2\right) \pm \frac{e_{\max} - e_{\min}}{2} \\
e &= \left(\hat{e}_1 + \hat{e}_2\right) \pm \frac{\left(\hat{e}_1 + \delta e_1\right) + \left(\hat{e}_2 + \delta e_2\right) - \left(\hat{e}_1 - \delta e_1\right) - \left(\hat{e}_2 - \delta e_2\right)}{2} \\
e &= \left(\hat{e}_1 + \hat{e}_2\right) \pm \frac{2(\delta e_1 + \delta e_2)}{2} \\
e &= \left(\hat{e}_1 + \hat{e}_2\right) \pm (\delta e_1 + \delta e_2)
\end{aligned}$$

เมื่อนำค่ามาบวกรวมกัน ความไม่แน่นอนของแต่ละค่าก็จะเอามารวมกันด้วยเช่นกัน

ในกรณีที่เรามีชิ้นงานที่มีความยาว $\hat{e}_1 \pm \delta e_1$ เราทำการใช้เครื่องมือตัดชิ้นส่วนโดยวัดจากขอบไปเป็นความยาว $\hat{e}_2 \pm \delta e_2$ เราจะหาความยาวของชิ้นงานที่โดนตัดได้ดังต่อไปนี้

$$e_{\max} = \left(\hat{e}_1 + \delta e_1\right) - \left(\hat{e}_2 - \delta e_2\right)$$
$$e_{\min} = \left(\hat{e}_1 - \delta e_1\right) - \left(\hat{e}_2 + \delta e_2\right)$$

หากเราเขียนอยู่ในรูป $\pm$ เราจะสามารถเขียนได้ดังต่อไปนี้

$$e = \left(\hat{e}_1 - \hat{e}_2\right) \pm \frac{p_{\max} - p_{\min}}{2}$$
$$e = \left(\hat{e}_1 - \hat{e}_2\right) \pm \frac{\left(\hat{e}_1 + \delta e_1\right) - \left(\hat{e}_2 - \delta e_2\right) - \left(\hat{e}_1 - \delta e_1\right) + \left(\hat{e}_2 + \delta e_2\right)}{2}$$
$$e = \left(\hat{e}_1 - \hat{e}_2\right) \pm \frac{2(\delta e_1 + \delta e_2)}{2}$$
$$e = \left(\hat{e}_1 - \hat{e}_2\right) \pm (\delta e_1 + \delta e_2)$$

ถึงแม้ว่าเราจะนำค่าสองค่ามาลบกัน แต่ความไม่แน่นอนจะถูกรวมกัน

การคูณและหารกันของความไม่แน่นอน

แทนที่เราจะวิเคราะห์ความไม่แน่นอนในเชิงสัมบูรณ์ เราสามารถวิเคราะห์ความไม่แน่นอนแบบสัมพัทธ์ $\delta^r e$ ได้ (relative uncertainty) ซึ่งสามารถคำนวณได้ตามความสัมพันธ์ดังต่อไปนี้

$$\delta^r e = \frac{\delta e}{|\hat{e}|}$$

โดยที่ $\hat{\delta} e$ คือความไม่แน่นอนสัมพัทธ์ของค่า e

สมมุติว่าแผนชิ้นงานถูกกำหนดมาให้มีรูปร่างเป็นสี่เหลี่ยมผืนผ้าที่มีความยาวเท่ากับ $\hat{e}_1 \pm \delta e_1$ และมีความยาวเท่ากับ $\hat{e}_2 \pm \delta e_2$ หากเราต้องการหาพื้นที่หรือผลคูณระหว่างความยาวและความกว้าง เราสามารถหาความไม่แน่นอนจากการรวมกันของความไม่แน่นอนสัมพัทธ์ซึ่งเขียนเป็นสมการได้ดังต่อไปนี้

$$\frac{\delta e}{|\hat{e}|} = \frac{\delta e_1}{|\hat{e}_1|} + \frac{\delta e_2}{|\hat{e}_2|}$$

$$\hat{e} \pm \delta e = \left(\hat{e}_1 \cdot \hat{e}_2\right) \pm |\hat{e}_1 \cdot \hat{e}_2| \cdot \frac{\delta e_1}{|\hat{e}_1|} + |\hat{e}_1 \cdot \hat{e}_2| \frac{\delta e_2}{|\hat{e}_2|}$$
$$\hat{e} \pm \delta e = \left(\hat{e}_1 \cdot \hat{e}_2\right) \pm \left|\frac{\hat{e}_1 \cdot \hat{e}_2}{\hat{e}_1}\right| \cdot \delta e_1 + \left|\frac{\hat{e}_1 \cdot \hat{e}_2}{\hat{e}_2}\right| \cdot \delta e_2$$
$$\hat{e} \pm \delta e = \left(\hat{e}_1 \cdot \hat{e}_2\right) \pm |\hat{e}_2| \cdot \delta e_1 + |\hat{e}_1| \cdot \delta e_2$$

สมมุติว่าเราต้องการคำนวณความชันของเส้นโดยกำหนดความต่างในแนวตั้งเป็น $e_1 \pm \delta e_1$ และ ความต่างในแนวนอนเป็น $e_2 \pm \delta e_2$ เราสามารถหาผลหารได้ในลักษณะเดียวกันกับผลคูณซึ่งเขียนเป็นสมการได้ดังต่อไปนี้

$$\frac{\delta e}{|\hat{e}|} = \frac{\delta e_1}{|\hat{e}_1|} + \frac{\delta e_2}{|\hat{e}_2|}$$

$$
\begin{aligned}
\widehat{e} \pm \delta e &= \left(\frac{\widehat{e}_1}{\widehat{e}_2}\right) \pm \left|\frac{\widehat{e}_1}{\widehat{e}_2}\right| \cdot \frac{\delta e_1}{|\widehat{e}_1|} + \left|\frac{\widehat{e}_1}{\widehat{e}_2}\right| \cdot \frac{\delta e_2}{|\widehat{e}_2|} \\
\widehat{e} \pm \delta e &= \left(\frac{\widehat{e}_1}{\widehat{e}_2}\right) \pm \left|\frac{\widehat{e}_1}{\widehat{e}_1 \cdot \widehat{e}_2}\right| \cdot \delta e_1 + \left|\frac{\widehat{e}_1}{\widehat{e}_2^2}\right| \cdot \delta e_2 \\
\widehat{e} \pm \delta e &= \left(\frac{\widehat{e}_1}{\widehat{e}_2}\right) \pm \frac{1}{\widehat{e}_2^2} \cdot \left(|\widehat{e}_2| \cdot \delta e_1 + |\widehat{e}_1| \cdot \delta e_2\right)
\end{aligned}
$$

ผลรวมเชิงเส้น (Linear Combination)

กำหนดให้ค่าที่ต้องคำนวณสามารถเขียนในรูปขงผลรวมเชิงเส้น (linear combination) ได้ดังต่อไปนี้

$$
\begin{aligned}
e &= a_1 \cdot \left(\widehat{e}_1 \pm \delta e_1\right) + a_2 \cdot \left(\widehat{e}_2 \pm \delta e_2\right) + \cdots \\
e &= \sum_{k=1}^{n} \left\{a_k \cdot \left(\widehat{e}_k \pm \delta e_k\right)\right\}
\end{aligned}
$$

โดยที่ a_k อาจจะเป็นบวกหรือลบก็ได้

เราสามารถหาค่าความคลาดเคลื่อนได้ดังต่อไปนี้

$$
e = \sum_{k=1}^{n} \left\{a_k \cdot \widehat{e}_k\right\} \pm \sum_{k=1}^{n} \left\{|a_k| \cdot \delta e_k\right\}
$$

การยกกำลังของความไม่แน่นอน

หากเรามีค่าที่มีความไม่แน่นอนที่โดนยกกำลัง เราสามารถหาความไม่แน่นอนได้โดยทำการคูณความไม่แน่นอนสัมพัทธ์เดิมด้วยเลขยกกำลัง

$$
\begin{aligned}
p &= (\widehat{e} \pm \delta e)^n \\
\frac{\delta p}{|\widehat{p}|} &= |n| \frac{\delta e}{|\widehat{e}|} \\
\delta p &= |n\widehat{e}^{n-1}| \delta e
\end{aligned}
$$

การประมาณความไม่แน่นอนด้วยอนุพันธ์

ในกรณีที่ความสัมพันธ์เป็นฟังก์ชันไม่เชิงเส้น เราสามารถประมาณความไม่แน่นอนจากนิยามของอนุพันธ์ กำหนดให้ความสัมพันธ์เป็นดังต่อไปนี้

$$
p = f(e)
$$

เราะสามารถประมาณความไม่แน่นอนได้โดยใช้นิยามดังต่อไปนี้

$$
\frac{\delta p}{\delta e} \approx \left|\frac{\mathrm{d}}{\mathrm{d}\,e}(f(e))\right|_{e=\widehat{e}}
$$

ดังนั้นเราจะได้ความสัมพันธ์ดังต่อไปนี้

$$
p \approx \widehat{e} \pm \left|\frac{\mathrm{d}}{\mathrm{d}\,e}(f(e))\right|_{e=\widehat{e}} \cdot \delta e
$$

เข้น

$$\cos(p) = \cos(\hat{e} \pm \delta e)$$
$$\cos(p) = \cos(\hat{e}) \pm \delta p = \cos(\hat{e}) \pm \left|\frac{\mathrm{d}}{\mathrm{d}\,e}\cos(e)\right|_{e=\hat{e}} \cdot \delta e$$
$$\cos(p) = \cos(\hat{e}) \pm |-\sin(\hat{e})| \cdot \delta e$$

การประยุกต์ใช้ความไม่แน่นอนของระบบ

เมื่อกล่าวถึงการจำลอง (simulation) คนส่วนมากมักจะคิดถึงการจำลองในอุดมคติที่มีค่าพารามิเตอร์ที่ชัดเจน แต่**ในโลกความเป็นจริง เราอาจจะไม่สามารถทราบค่าที่ชัดเจนของพารามิเตอร์เหล่านี้ เพื่อที่จะจำลองให้มีลักษณะที่ครอบคลุมหรือใกล้เคียงกับการทดลองจริง เราสามารถประยุกต์ใช้ความไม่แน่นอนนี้ได้** หากเราสามารถกำหนดช่วงของพารามิเตอร์เหล่านี้ได้ เราจะสามารถจำลองเป็นจำนวนหลายครั้งได้โดยเปลี่ยนแปลงพารามิเตอร์ในช่วงความไม่แน่นอนที่คำนวณมา

ตัวอย่างที่ 1 : การจำลองพลวัต

ตัวอย่างนี้เป็นการทดลองเพื่อวัดหาเวลาที่ทำให้วัตถุหยุดจากความเร็วเชิงมุมที่คงที่ ω^*

วัตถุที่ถูกทำให้หมุนเป็นวัตถุทรงปริซึมสี่เหลี่ยมผืนผ้าซึ่งสามารถหาความเฉื่อยในแนวแกนหมุนตามสมการดังต่อไปนี้

$$J = \frac{1}{12}m(a^2 + b^2) + ml^2$$

โดยที่

a และ b เป็นความกว้างและความยาวของวัตถุ $[m]$

l เป็นระยะห่างระหว่างจุดศูนย์กลางมวลและแกนในการหมุน $[m]$

m เป็นมวลของวัตถุ $[\mathrm{kg}]$

นอกจากนี้ ระบบที่โดนหมุนได้เชื่อมต่อกับตลับลูกปืนที่มีความฝืดเชิงมุมเท่ากับ B $\left[\frac{\mathrm{N \cdot m \cdot s}}{\mathrm{rad}}\right]$

และมีพลวัตตามสมการดังต่อไปนี้

$$J \cdot \frac{\mathrm{d}}{\mathrm{d}\,t}(\omega) = -B \cdot \omega$$
$$\omega(0) = \omega^*$$

เราจะนิยามระยะเวลาในการหยุดของวัตถุ T_s ให้เป็นระยะเวลาที่ความเร็วของวัตถุมีความเร็วเหลือแค่ 2% ของความเร็วเริ่มต้น

$$\omega(T_s) = 0.02 \cdot \omega^*$$

นอกจากลักษณะและความสัมพันธ์ของระบบและพารามิเตอร์ที่บอกมา ค่าต่างๆและความไม่แน่นอนถูกกำหนดมาดังนี้

$$\begin{bmatrix} \text{พารามิเตอร์} & \text{ค่าของพารามิเตอร์} & \text{หน่วย} \\ \omega^* & 60 \pm 2.5 & \mathrm{rpm} \\ B & 0.067 \pm 0.034 & \dfrac{\mathrm{N\cdot m\cdot s}}{\mathrm{rad}} \\ m & 1.343 \pm 0.001 & \mathrm{kg} \\ a & 0.35 \pm 0.005 & \mathrm{m} \\ b & 0.25 \pm 0.005 & \mathrm{m} \\ l & 0.05 \pm 0.005 & \mathrm{m} \end{bmatrix}$$

หนึ่งในวิธีที่เราสามารถจำลองระบบให้ครอบคลุมได้คือการกำหนดให้พารามเตอร์มีค่าที่เป็นได้

```
N = 11;
w_0 = 60+2.5*linspace(-1,1,N);
B = 0.067+0.034*linspace(-1,1,N);
m = 1.343+0.001*linspace(-1,1,N);
a = 0.35+0.005*linspace(-1,1,N);
b = 0.25+0.005*linspace(-1,1,N);
l = 0.05+0.005*linspace(-1,1,N);
```

หากเรากำหนดค่าพารามิเตอร์ในรูปแบบดังกล่าว เราจะต้องจำลองทั้งหมด 11^6 **หรือมากกว่า 1.7 ล้านครั้ง** แต่ถ้าหากเราหาความไม่แน่นอนและรวมพารามิเตอร์บางตัวได้ เราจะสามารถลดการคำนวณได้อีก เราทำการย้ายพารามิเตอร์ของระบบให้เป็นดังต่อไปนี้

$$\frac{\mathrm{d}}{\mathrm{d}\,t}(\omega) = -\frac{1}{\tau}\cdot\omega$$

$$\tau = \frac{J}{B}$$

เราคำนวณค่าความไม่แน่นอนของพารามิเตอร์ k ได้ดังต่อไปนี้

$$\delta\tau = \hat{\tau}\cdot\left(\frac{\delta J}{\hat{J}} + \frac{\delta B}{\hat{B}}\right)$$

$$\delta J = \frac{1}{12}\delta\{m(a^2+b^2)\} + \delta\{ml^2\}$$

$$\delta J = \frac{1}{12}\cdot\left[\hat{m}\cdot\delta\{(a^2+b^2)\} + \left(\hat{a}^2+\hat{b}^2\right)\cdot\delta m\right] + \left[\hat{m}\cdot\delta\{l^2\} + \hat{l}^2\cdot\delta m\right]$$

$$\delta J = \frac{1}{12}\cdot\left[\hat{m}\cdot\left(2\hat{a}\cdot\delta a + 2\hat{b}\cdot\delta b\right) + \left(\hat{a}^2+\hat{b}^2\right)\cdot\delta m\right] + \left[\hat{m}\cdot 2\hat{l}\cdot\delta l + \hat{l}^2\cdot\delta m\right]$$

$$\delta J = \left(\frac{\hat{m}\hat{a}}{6}\right)\cdot\delta a + \left(\frac{\hat{m}\hat{b}}{6}\right)\cdot\delta b + \left(\frac{\hat{a}^2+\hat{b}^2}{12} + \hat{l}^2\right)\cdot\delta m + \left(2\hat{m}\hat{l}\right)\cdot\delta l$$

```
w_0_e = 60;
B_e = 0.067;
m_e = 1.343;
a_e = 0.35;
b_e = 0.25;
l_e = 0.05;

dw = 2.5;
dB = 0.034;
dm = 0.001;
da = 0.005;
```

```
db = 0.005;
dl = 0.005;

dJ = (m_e*a_e/6)*da+(m_e*b_e/6)*db+((a_e^2+b_e^2)/12+l_e^2)*dm+(2*m_e*l_e)*dl;
J_e = (m_e*(a_e^2+b_e^2))/12+m_e*l_e^2;
tau_e = J_e/B_e;
dtau = tau_e*(dJ/J_e+dB/B_e);
```

ผลที่ได้คือค่า k ที่มีค่าเท่ากับ $2.7845 \pm 1.5705\ [s^{-1}]$

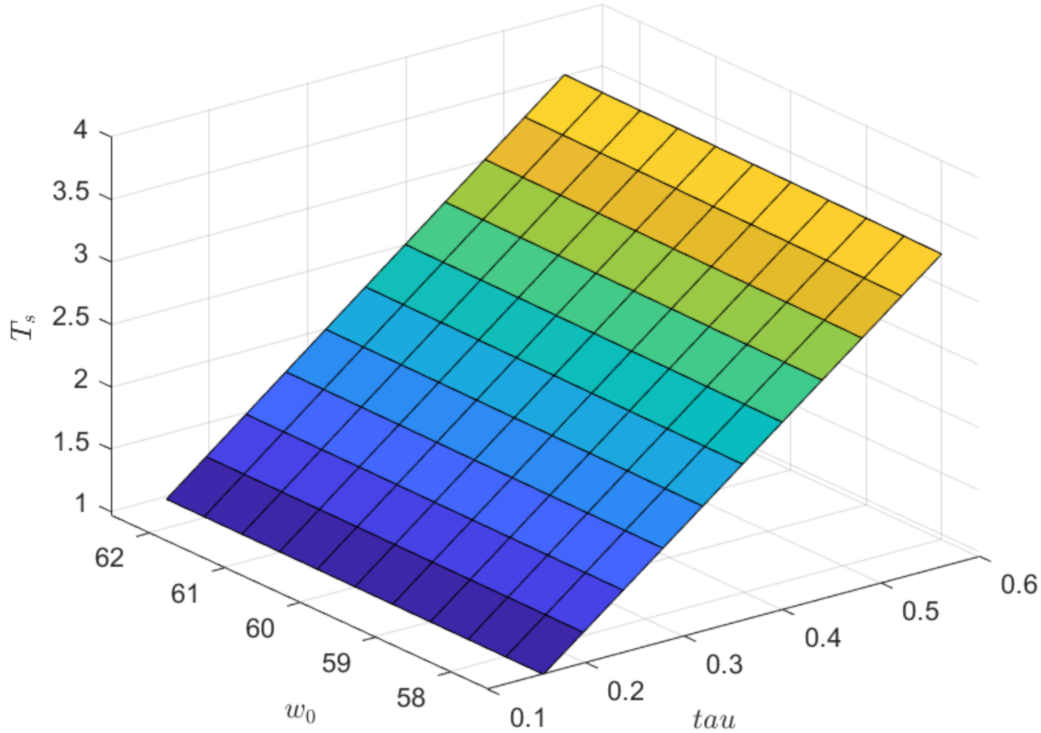
$$\begin{bmatrix} \text{พารามิเตอร์} & \text{ค่าของพารามิเตอร์} & \text{หน่วย} \\ \omega^* & 60 \pm 2.5 & \text{rpm} \\ k & 0.3591 \pm 0.2026 & \text{s} \end{bmatrix}$$

```
N = 11;
w_0 = 60+2.5*linspace(-1,1,N);
tau = 0.3591+0.2026*linspace(-1,1,N);
```

เราจะสังเกตได้ว่า **พารามิเตอร์ที่ไม่แน่นอนนั้นถูกลดจำนวนจาก 6 ตัวเหลือแค่ 2 ตัว ทำให้ตำนวนครั้งในการจำลองลดลงเหลือ** 121 **ครั้ง** หากการจำลองหนึ่งครั้งต้องใช้เวลา $0.1\ [s]$ **เดิมทีเราต้องใช้เวลา 49 ชั่วโมงในการจำลอง แต่การลดจำนวนตัวแปรทำให้การจำลองลดเวลาเหลือแค่ 12 วินาที**

```
N = 11;
w_e = 60;
w_0 = 60+2.5*linspace(-1,1,N);
tau_e = 0.3591;
tau = 0.3591+0.2026*linspace(-1,1,N);
% at2percent = @(t,y,w_e)deal(y==0.002*w_e,1,-1);
opt = odeset('Events',@(t,y)at2percent(t,y,w_e));
t_max = 10*tau_e;
tspan = [0 t_max];
T_s = zeros(N,N);
for i = 1:N
    for j = 1:N
        [t,y] = ode45(@(t,y)-y/tau(j),tspan,w_0(i),opt);
        T_s(i,j) = t(end);
    end
end

[T,W] = meshgrid(tau,w_0);
clf
ax = axes;
surf(ax,T,W,T_s)
xlabel('$tau$','Interpreter',"latex")
ylabel('$w_0$','Interpreter',"latex")
zlabel('$T_s$','Interpreter',"latex")
```




```
T_mean = mean(T_s,'all');
T_max = max(T_s,[],'all');
T_min = min(T_s,[],'all');
```

จากจำลอง 121 ครั้ง ค่าเฉลี่ยที่ $2.2319\ [\mathrm{s}]$ ค่าที่มากที่สุดที่ $3.5139\ [\mathrm{s}]$ และค่าที่น้อยที่สุดที่ $0.9661\ [\mathrm{s}]$ ซึ่งเราเขียนเป็นความไม่แน่นอนได้ที่ $2.2319 \pm 1.2820\ [\mathrm{s}]$ สาเหตุที่ความไม่แน่นอนสัมพันธ์อยู่ที่ $57.44\ \%$ ก็เพราะว่าความไม่แน่นอนของค่าความฝืดมีความไม่แน่นอนสัมพัทธ์ที่สูงอยู่แล้ว (50.75%) ดังนั้นผลจากการจำลองเลยมีความไม่แน่นอนที่สูงตามเช่นกัน

```
function [condition,isTerminal,direction] = at2percent(t,y,w_e)
condition = y<=0.002*w_e;
isTerminal = 1;
direction = 1;
end
```

ความน่าจะเป็นและสัญญาณสุ่ม (Probability & Random Signal)

ตอนที่ 2 : ตัวแปรสุ่ม (1/2) (Random Variable)

จะอย่างไรก็ตามการอธิบายความไม่แน่นอนในลักษณะบวกลบนี้มีข้อเสียในเรื่องของการหาแนวโน้มและความสัมพันธ์เนื่องจาก**การระบุความไม่แน่นอนแบบบวกลบไม่ได้ให้ข้อมูลการแจกแจงของค่ามาแต่อย่างใด** ถึงแม้ค่าจะมีช่วงที่กำหนดมา แต่ค่าเฉลี่ยจริงๆนั้นอาจจะไม่ได้อยู่ที่กึ่งกลางของช่วงก็เป็นได้ อีกทั้งความแปรปรวนของค่าอาจจะมีมากหรือน้อยที่แตกต่างกันก็เป็นได้ ดังนั้น ในหลายครั้ง เราจะใช้ความน่าจะเป็น (probability) ในการอธิบายค่าที่เป็นไปได้ของตัวแปร

ตัวแปรที่มีความไม่แน่นอนจะถูกเรียกว่า ตัวแปรสุ่ม (random variable)ซึ่ง เราสามารถอธิบายค่าที่เป็นไปได้ของตัวแปรสุ่มนี้ผ่านความน่าจะเป็น (probability) ซึ่งตัวแปรนี้สามารถมีค่าเป็นได้ 2 ประเภทหลัก

ตัวแปรสุ่มแบบไม่ต่อเนื่อง (discrete random variable)

ประเภทแรกคือ**ตัวแปรสุ่มแบบไม่ต่อเนื่อง (discrete random variable) ซึ่งเป็นตัวแปรที่มีค่าที่เป็นไปได้อยู่ในเซ็ตที่นับได้** ยกตัวอย่างการโยนเหรียญหัวก้อย หากเรากำหนดให้ Xเป็นตัวแปรสุ่มที่บ่งบอกว่าเหรียญเป็นหัวหรือเป็นก้อย โดยกำหนดให้ เมื่อ $X=1$ เป็นเหตุการณ์ที่เหรียญออกเป็นหัว ส่วน $X=0$ เป็นเหตุการณ์ที่เหรียญออกเป็นก้อย เราจะสามารถเขียนความน่าจะเป็น $\mathrm{P}(\cdot)$ ได้ดังต่อไปนี้

$$\mathrm{P}(X=1)=0.5$$
$$\mathrm{P}(X=0)=0.5$$

สมการการทั้งสองอธิบายถึงความน่าจะเป็นที่เหรียญจะออกหัวหรือก้อยนั้นเป็น 0.5 ทั้งคู่ (50%)

ความน่าจะเป็นนั้นจะเอาไว้ใช้อธิบายเหตุการณ์ (event) เสมอ ในการอธิบายอย่างเป็นทางการ **เราจำเป็นต้องแปลงความหมายของเหตุการณ์ให้อยู่ในรูปของสมการหรืออสมการ** ยกตัวอย่างเช่นการทอยลูกเต๋าหกด้าน หากต้องการหาความน่าจะเป็นที่ลูกเต๋าออกด้านที่มีค่าน้อยกว่า 4 เราต้องกำหนดให้ X เป็นค่าของด้านลูกเต๋าที่ออก และเหตุการณ์ที่ลูกเต๋าออกด้านที่มีค่าน้อยกว่า 4 จะสามารถถูกเขียนเป็นอสมการได้คือ $X<4$ ดังนั้น ความน่าจะเป็นสามารถเขียนได้เป็นสมการต่อไปนี้

$$\mathrm{P}(X<4)=0.5$$

<u>เราจะไม่เขียนว่า $\mathrm{P}(X)$ เนื่องจาก X ไม่ใช่เหตุการณ์</u>

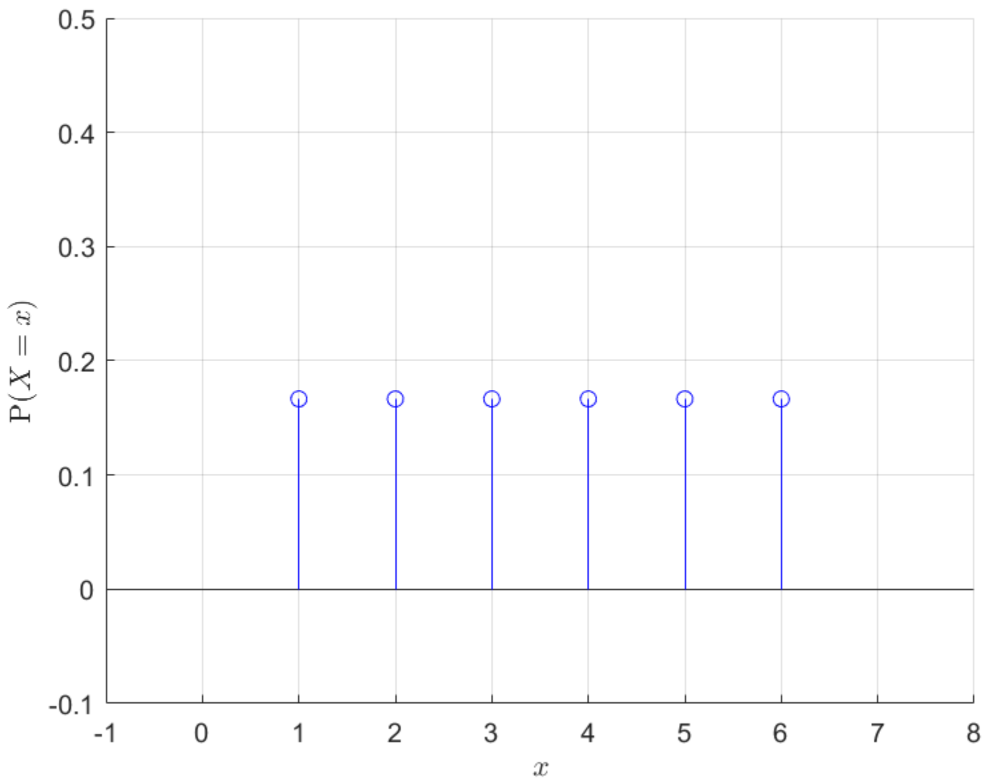
ตัวอย่างที่ 1 : การทอยลูกเต๋า

กำหนดให้ลูกเต๋าหกด้านมีความน่าจะเป็นที่เท่าๆกันที่จะออกด้านใดด้านหนึ่ง ดังนั้น เราสามารถเขียนความสัมพันธ์ของความน่าจะเป็นได้ดังต่อไปนี้

$$\mathrm{P}(X=x)=\frac{1}{6} \quad \text{โดยที่ } x\in\{1,2,3,4,5,6\}$$

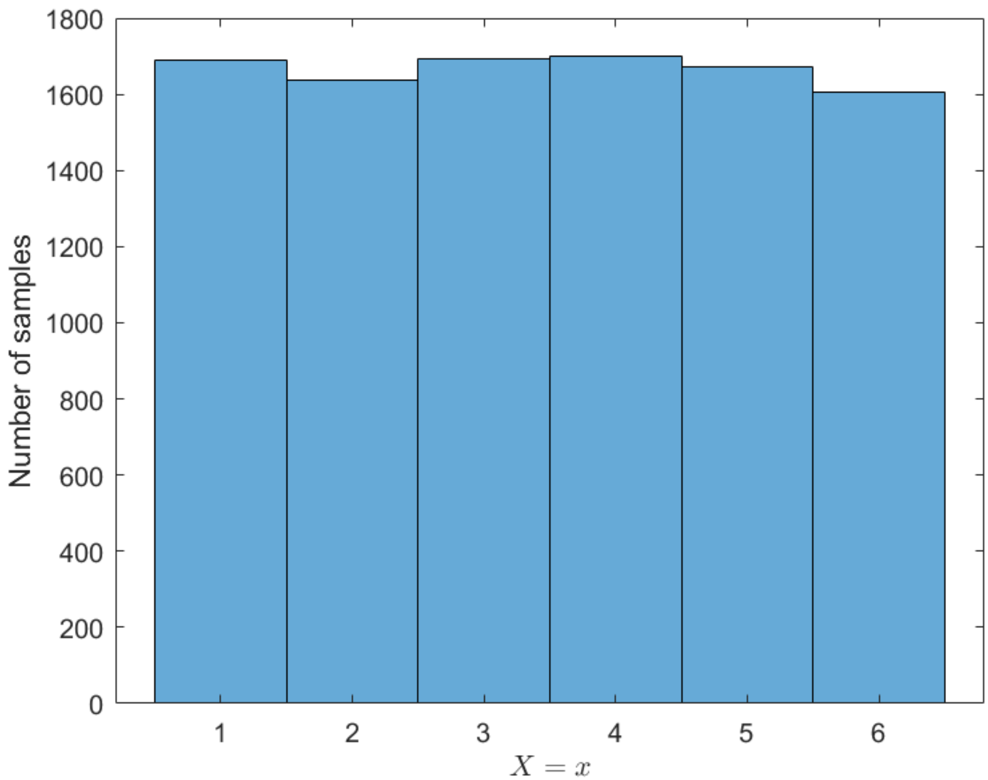
เราสามารถกราฟความสัมพันธ์ระหว่างค่าของ X และความน่าจะเป็นได้ดังต่อไปนี้

```
clf;
ax = axes;
grid(ax,'on')
hold(ax,'on')
x = 1:6;
stem(ax,x,(1/6)*ones(size(x)),'b')
axis(ax,[-1 8 -0.1 0.5])
xlabel('$x$','Interpreter',"latex")
ylabel('$\mathrm{P}(X=x)$','Interpreter',"latex")
```



การแจกแจงที่ความน่าจะเป็นของทุกเหตุการณ์เท่ากันถูกเรียกว่า การแจกแจงแบบยูนิฟอร์ม **(uniform distribution)** ซึ่งเราสามารถเขียนเป็นโปรแกรมสุ่มโดยใช้ฟังก์ชัน **randi** ได้ดังต่อไปนี้

```
imax = 6;
imin = 1;
sz = [1 10000]; % one row 100 columns
X = randi(imax-imin+1,sz)+imin-1;
clf
ax = axes;
histogram(ax,X)
xlabel('$X=x$','Interpreter',"latex")
ylabel('Number of samples')
```



หากโจทย์เปลี่ยนเป็นการทอยลูกเต๋า 2 อันและนับผลรวมจากแต้มของหน้าลูกเต๋าทั้งสอง เราจะได้ความน่าจะเป็นของแต่ละผลรวมดังต่อไปนี้

$$\mathrm{P}(X = 2) = \mathrm{P}(X = 12) = \frac{1}{36}$$

$$\mathrm{P}(X = 3) = \mathrm{P}(X = 11) = \frac{2}{36}$$

$$\mathrm{P}(X = 4) = \mathrm{P}(X = 10) = \frac{3}{36}$$

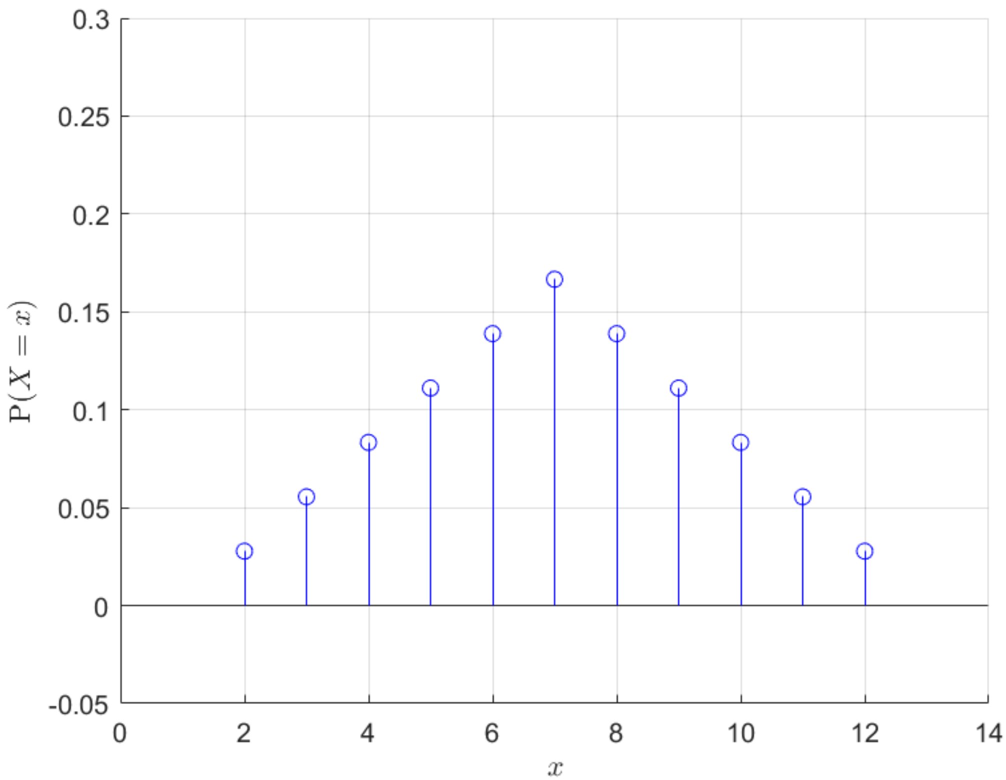
$$\mathrm{P}(X = 5) = \mathrm{P}(X = 9) = \frac{4}{36}$$

$$\mathrm{P}(X = 6) = \mathrm{P}(X = 8) = \frac{5}{36}$$

$$\mathrm{P}(X = 7) = \frac{6}{36}$$

โดยที่ X คือผลรวมของหน้าลูกเต๋าที่ออก ซึ่งสามารถกราฟความสัมพันธ์ได้ดังต่อไปนี้

```
clf;
ax = axes;
grid(ax,'on')
hold(ax,'on')
x = 2:12;
stem(ax,x,(6-abs(x-7))/36,'b')
axis(ax,[0 14 -0.05 0.3])
xlabel('$x$','Interpreter',"latex")
ylabel('$\mathrm{P}(X=x)$','Interpreter',"latex")
```



ในบ่อยครั้ง เหตุการณ์อาจจะมีจำนวนมากซึ่งการเขียนสมการเป็นทีละเหตุการณ์อาจไม่ใช่วิธีที่สะดวกที่สุด เราสามารถใช้ฟังก์ชันในการอธิบายการแจกแจงของความน่าจะเป็น หรือ ฟังก์ชันการแจกแจงความน่าจะเป็น $f_X(\cdot)$ (probability density function/ pdf) สำหรับตัวอย่างการทอยลูกเต๋าสองอัน เราสามารถเขียนฟังก์ชันการแจกแจงความน่าจะเป็นได้ดังต่อไปนี้

$$f_X(x) = (x-1)\cdot\theta(x-2) - 2(x-7)\cdot\theta(x-7) + (x-13)\cdot\theta(x-13)$$

โดยที่ x คือค่าของผลรวมหน้าลูกเต๋า (จำนวนเต็ม)

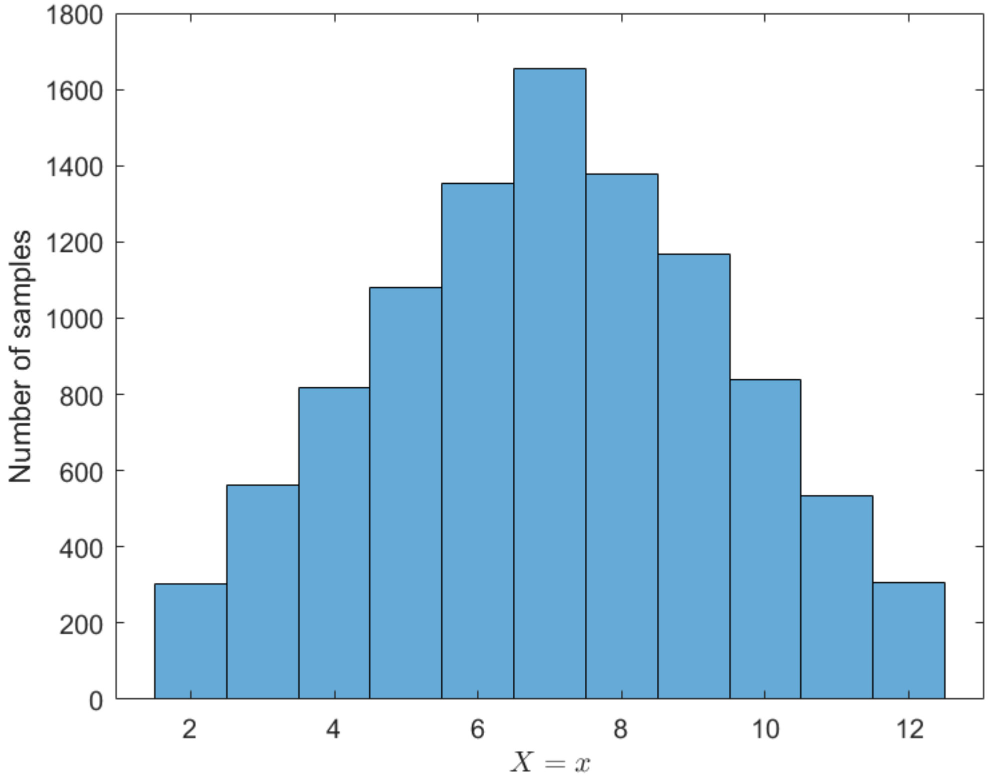
ไม่ว่าความน่าจะเป็นจะมีการแจกแจงอย่างไร เราสามารถเขียนความน่าจะเป็นของแต่ละจำนวนเต็มให้อยู่ในรูปของเวกเตอร์ที่มีจำนวนจำกัดได้ โดยการกำหนดตัวแปรที่มีจำนวนจำกัดและใช้ผ่านฟังก์ชันเช่น

$$f_X(x) = \frac{6-|x-7|}{36} \text{ โดยที่ } x \in \{2,3,\cdots,12\}$$

ถ้าหากเราเขียนความน่าจะเป็นในรูปแบบที่จำกัดได้ เราก็สามารถเขียนโปรแกรมในการสุ่มตามการแจกแจงใดๆได้ ซึ่งเราจะใช้ฟังก์ชัน randsample (จำเป็นต้องใช้ Statistics & Machine Learning Toolbox) ได้ดังต่อไปนี้

```
x = 2:12;
pdf = (6-abs(x-7))/36;
sz = 10000;
X = randsample(x,sz,true,pdf);

clf
ax = axes;
histogram(ax,X)
xlabel('$X=x$','Interpreter',"latex")
ylabel('Number of samples')
```



หนึ่งในคุณสมบัติของฟังก์ชันการแจกแจงความน่าจะเป็นสำหรับตัวแปรที่ไม่ต่อเนื่องคือผลรวมของความน่าจะเป็นของทุกเหตุการณ์ต้องเท่ากับ 1 หรือเขียนเป็นสมการได้ดังต่อไปนี้

$$\sum_x f_X(x) = 1$$

อีกหนึ่งในคุณสมบัติของฟังก์ชันการแจกแจงคือฟังก์ชันต้องไม่เป็นลบ

$$0 \leq f_X(x) \leq 1$$

ความน่าจะเป็นและสัญญาณสุ่ม (Probability & Random Signal)

ตอนที่ 3 : ตัวแปรสุ่ม (2/2) (Random Variable)

ตัวแปรสุ่มแบบต่อเนื่อง (continuous random variable)

ตัวแปรสุ่มประเภทที่สองคือ**ตัวแปรสุ่มแบบต่อเนื่อง (continuous random variable) ซึ่งเป็นตัวแปรที่มีค่าที่เป็นไปได้อยู่ในเซ็ตที่นับไม่ได้** ยกตัวอย่างเช่น เวลาที่รถไฟจะมาถึงสถานี หรือ ความยาวของเส้นผ่านศูนย์กลางของเพลาที่ถูกกลึง ค่าเหล่านี้ล้วนมีค่าที่เป็นไปได้ที่ต่อเนื่อง (conitnuous)

หากเรากำหนดให้ระยะเวลาที่รถไฟจะมาถึงสถานีเป็น X ถึงแม้ว่ารถไฟจะถูกจัดตารางมาให้ถึงทุกๆ 10 นาที ความน่าจะเป็นที่ระยะเวลานั้นจะเท่ากับ 10 นาที จะมีค่าเท่ากับ 0 ในทางทฤษฎี เหตุการณ์ที่ระยะเวลาจะเท่ากับ 10 นาที มีแค่ครั้งเดียวเทียบกับเหตุการณ์อื่นๆที่ระยะเวลาเท่ากับค่าอื่นเช่น 10.1, 10.01, 10.001, 10.0001, 10.00000001, 9.9, 9.99, 9.9999 **เรานั้นสามารถเขียนเหตุการณ์ที่เกิดขึ้นนั้นได้เป็นจำนวนครั้งไม่ถ้วนเนื่องจากค่าของตัวแปรเป็นค่าตอเนื่อง** ดังนั้น ความน่าจะเป็นของเหตุการณ์ที่ตัวแปรสุ่มมีค่าเท่ากับค่าคงที่ค่าหนึ่งนั้นจะเป็น 0 เสมอ

$$\mathrm{P}(X=0)=0$$

หรือกล่าวได้อีกอย่างคือ **เราไม่ควรใช้สมการในการอธิบายเหตุการณ์สำหรับตัวแปรสุ่มแบบต่อเนื่อง**

ตัวอย่างที่ 1 : การกลึงเพลา

โรงงานมีเครื่องกลึงอยู่ 2 เครื่อง ซึ่งแต่ละเครื่องถูกตั้งค่าไว้สำหรับกลึงเพลาให้มีเส้นผ่านศูนย์กลาง ขนาด $4\,[\mathrm{cm}]$

กำหนดให้ D_A และ D_B เป็นขนาดของเส้นผ่านศูนย์กลางของเพลาที่สร้างโดยเครื่องกลึง A และเครื่องกลึง B ตามลำดับ

หากเรากำหนดความเผื่อ (tolerance) ในการผลิตไว้ที่ $0.01\,[\mathrm{cm}]$ เราสามารถเปรียบเทียบสมรรถนะของเครื่องกลึงแต่ละเครื่องได้ดังนี้

$$\mathrm{P}(|D_A-4|\le 0.01)=0.6827$$
$$\mathrm{P}(|D_B-4|\le 0.01)=0.9545$$

จากตัวอย่างที่กำหนด ทำให้เราสรุปได้ว่า เครื่องกลึง B มีสมรรถนะที่ดีกว่า (จาก หนึ่งร้อยชิ้นที่ผลิต เพลาจากเครื่องกลึง A จำนวน 68 ชิ้น ผ่านตามเกณฑ์ที่กำหนด แต่เพลาจากเครื่องกลึง B จำนวน 95 ชิ้นผ่านตามเกณฑ์ที่กำหนด)

การแจกแจงของความน่าจะเป็น (probability distribution)

เหตุการณ์ที่เกี่ยวข้องกับตัวแปรสุ่มแบบต่อเนื่องจะถูกอธิบายโดยใช้อสมการ หรือ ช่วงที่เป็นไปได้ของค่า เช่น

$$\mathrm{P}(X\le 1)\ ,\ \mathrm{P}(-0.3\le X\le 0.1)\ ,\ \mathrm{P}(|x-2.5|\le 0.1)$$

เช่นดียวตัวแปรสุ่มแบบไม่ต่อเนื่อง การแจกแจงของความน่าจะเป็นของตัวแปรต่อเนื่องสามารถเขียนอยู่ในรูปของฟังก์ชันได้ ซึ่งต้องมีความสัมพันธ์ดังต่อไปนี้

$$\int_{-\infty}^{\infty}\{f_X(x)\}\mathrm{d}\,x=1$$
$$0\le f_X(x)\le 1,\quad 0\ \forall x$$

เรากำหนดให้ความสัมพันธ์ระหว่าง ความน่าจะเป็น และการแจกแจงของความน่าจะเป็นดังต่อไปนี้

$$\mathrm{P}(X\le k)=\int_{-\infty}^{k}\{f_X(x)\}\mathrm{d}\,x$$

ความน่าจะเป็นที่ตัวแปรสุ่มจะมีค่าน้อยกว่าหรือเท่ากับค่า k มีค่าเท่ากับการหาพื้นที่ใต้กราฟของฟังก์ชันการแจกแจงความน่าจะเป็นจากลบอนันต์ถึงค่า k

จากความสัมพันธ์ดังกล่าว เราสามารถคำนวณหาความน่าจะเป็นของเหตุการณ์ที่ถูกเขียนในรูปของอสมการดังต่อไปนี้

$$\begin{aligned}\mathrm{P}(X \geq k) &= 1 - \mathrm{P}(X \leq k) = 1 - \int_{-\infty}^{k} \{f_X(x)\}\mathrm{d}\,x \\ &= \int_{-\infty}^{\infty} \{f_X(x)\}\mathrm{d}\,x - \int_{-\infty}^{k} \{f_X(x)\}\mathrm{d}\,x \\ &= \int_{k}^{\infty} \{f_X(x)\}\mathrm{d}\,x\end{aligned}$$

$$\begin{aligned}&\mathrm{P}(|X-\mu| \leq k) = \mathrm{P}(\mu - k \leq X \leq \mu + k) = 1 - \mathrm{P}(X \leq \mu - k) - \mathrm{P}(X \geq \mu + k) \\ &= \int_{-\infty}^{\infty} \{f_X(x)\}\mathrm{d}\,x - \int_{-\infty}^{\mu-k} \{f_X(x)\}\mathrm{d}\,x - \int_{\mu+k}^{\infty} \{f_X(x)\}\mathrm{d}\,x \\ &= \int_{\mu-k}^{\mu+k} \{f_X(x)\}\mathrm{d}\,x\end{aligned}$$

ดังนั้น เราสามารถอธิบายการแจกแจงโดยใช้ฟังก์ชันได้ตราบใดที่ฟังก์ชันนั้นมีคุณสมบัติต่อไปนี้

$$\begin{aligned}&\int_{-\infty}^{\infty} \{f_X(x)\}\mathrm{d}\,x = 1 \\ &f_X(x) \geq 0 \;\; \forall x\end{aligned}$$

การแจกแจงแบบยูนิฟอร์ม (Uniform Distribution)

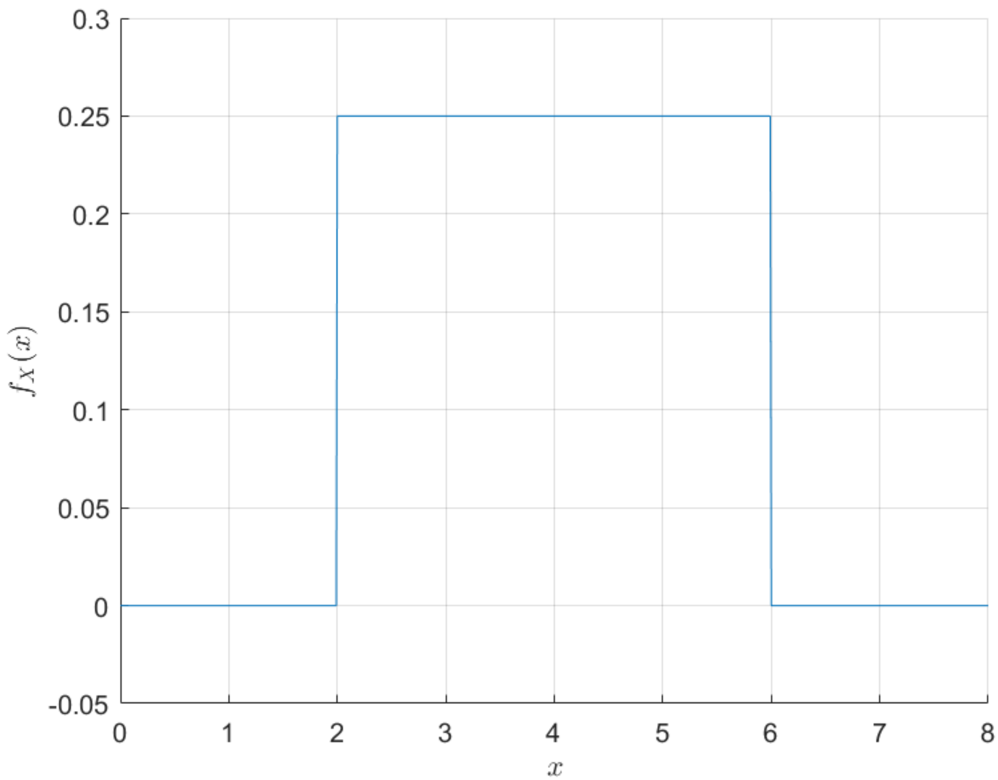
ในกรณีที่เหตุการณ์นั้นเกิดขึ้นอยู่ในช่วงที่จำกัด

$$X \in [x_{\min}, x_{\max}]$$

และมีการแจกแจงความน่าจะเป็นอย่างเท่าๆกันทุกเหตุการณ์ เราสามารถเขียนฟังก์ชันการแจกแจงความน่าจะเป็นได้ดังต่อไปนี้

$$f_X(x) = \frac{\theta(x - x_{\min}) - \theta(x - x_{\max})}{x_{\max} - x_{\min}}$$

```
clf
ax = axes;
hold(ax,'on')
grid(ax,'on')
x_max = 6;
x_min = 2;
x = 0:0.01:8;
plot(ax,x,((x>=x_min)-(x>=x_max))/(x_max-x_min))
xlabel('$x$','Interpreter',"latex");
ylabel('$f_X(x)$','Interpreter',"latex")
axis(ax,[0 8 -0.05 0.3])
```



เราสามารถนำแนวคิดนี้ไปเขียนเป็นโปรแกรมที่ใช้ในการสุ่มตัวเลขที่เป็นค่าต่อเนื่องแต่อยู่ในช่วงได้ดังต่อไปนี้

```
x_max = 6;
x_min = 2;
sz = [1,100];
X = (x_max-x_min)*rand(sz)+x_min;
```

การสุ่มแบบดังกล่าวเหมาะสำหรับตัวแปรสุ่มที่เราไม่ทราบค่าแน่ชัดแต่เราสามารถหาช่วงของค่านั้นได้และทุกค่าที่อยู่ในช่วงมีโอกาสที่จะเป็นได้เท่าๆกัน

การแจกแจงแบบปรกติ (Normal (Gaussian) Distribution)

ในหลายครั้ง การแจกจางของความน่าจะเป็นไม่ได้มีค่าเท่ากันหมดในทั้งช่วง การแจกแจงของความน่าจะเป็นนั้นอาจจะเกาะกลุ่มใกล้ค่าค่าหนึ่งมากกว่าค่าอื่นๆในช่วง **หนึ่งในการแจกแจงความน่าจะเป็นที่พบเห็นบ่อยในการิวเคราะห์เชิงสถิติและงานวิจัยได้แก่การแจกแจงแบบปรกติ** หรือ การแจกแจงแบบเกาส์เซียน (normal/Guassian distribution) ซึ่งสามารถเขียนเป็นฟังก์ชันได้ดังต่อไปนี้

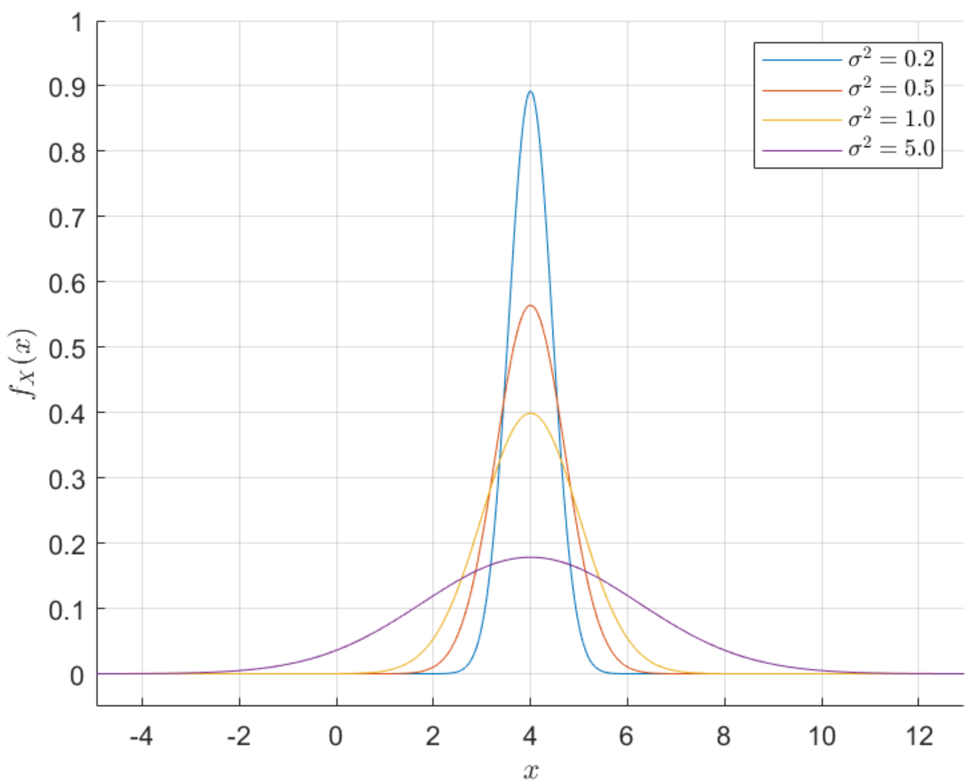
$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2}$$

โดยที่

μ_x และ σ_X เป็นพารามิเตอร์ของฟังก์ชันนี้ (ซึ่งมีบทบาทเกี่ยวข้องกับค่าคาดหวังและค่าแปรปรวน)

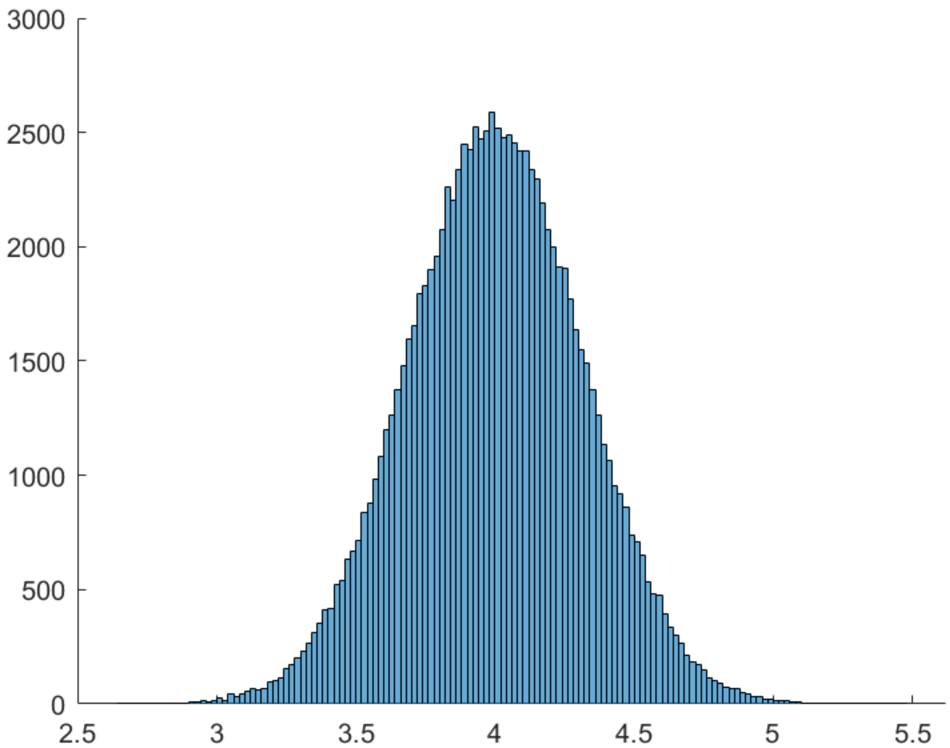
```
clf
ax = axes;
hold(ax,'on')
```

```
grid(ax,'on')
mu = 4;
sigma_sqr = [0.2 0.5 1 5];
sigma = sqrt(sigma_sqr);
lab = cell(1,numel(sigma));
 x = (-4*max(sigma):0.01:4*max(sigma))+mu;
for i = 1:numel(sigma)
    s = sigma(i);
    plot(ax,x,exp(-0.5*((x-mu)/s).^2)/sqrt(2*pi*s^2))
    xlabel('$x$','Interpreter',"latex");
    ylabel('$f_X(x)$','Interpreter',"latex")
    axis(ax,[min(x) max(x) -0.05 1]);
    lab{i} = sprintf('$\\sigma^2 = %.1f$',sigma_sqr(i));
end
legend(lab,'Interpreter',"latex")
```



ค่า σ_X ของฟังก์ชันบ่งบอกถึงความแปรปรวนของค่า X ในบริเวณรอบค่า μ_X ยิ่งค่าของ σ_X มีค่าน้อย ความน่าจะเป็นที่ X อยู่บริเวณรอบๆ μ_X จะยิ่งมีมากขึ้น **การแจกแจงความน่าจะเป็นแบบดังกล่าวมีประโยชน์ถ้าเราทราบถึงค่าคาดหวังและค่าแปรปรวนของตัวแปรสุ่ม ซึ่งเราสามารถนำมาเขียนเป็นโปรแกรมในการสุ่มได้ดังนี้**

```
sigma_sqr = 0.1;
mu = 4;
X = sqrt(sigma_sqr)*randn([1,100000])+mu;
clf;
ax = axes;
hold(ax,'on')
histogram(ax,X)
```



การแจกแจงของความน่าจะเป็นแบบปรกติเป็นการแจกแจงที่ได้รับความนิยม เนื่องจากมีความเกี่ยวข้องกับ ทฤษฎี Central Limit Theorem และพฤติกรรมหรือพลวัตต่างๆ ตามธรรมชาติล้วนแต่ให้ผลลัพธ์ทางสถิติที่มีลักษณะเป็นระฆังคว่ำหรือการแจกแจงแบบปรกตินั่นเอง

การแจกแจงแบบปรกติด้วยล็อก (Log-Normal Distribution)

สมมุติว่าเรากำหนดอัตราเร็วเชิงเส้น (linear speed) ของหุ่นยนต์ที่ค่า V แต่ในโลกความจริง อัตราเร็วของหุ่นยนต์นั้นมีความไม่แน่นอนอยู่ เราสามารถจำลองความน่าจะเป็นของอัตราเร็วด้วยการเลือกลักษณะการแจกแจง อย่างไรก็ตาม ค่าของอัตราเร็วนั้นมีค่าเป็นที่มากกว่าหรือเท่ากับ 0 ดังนั้นเราไม่สามารถใช้การแจกแจงแบบปรกติได้

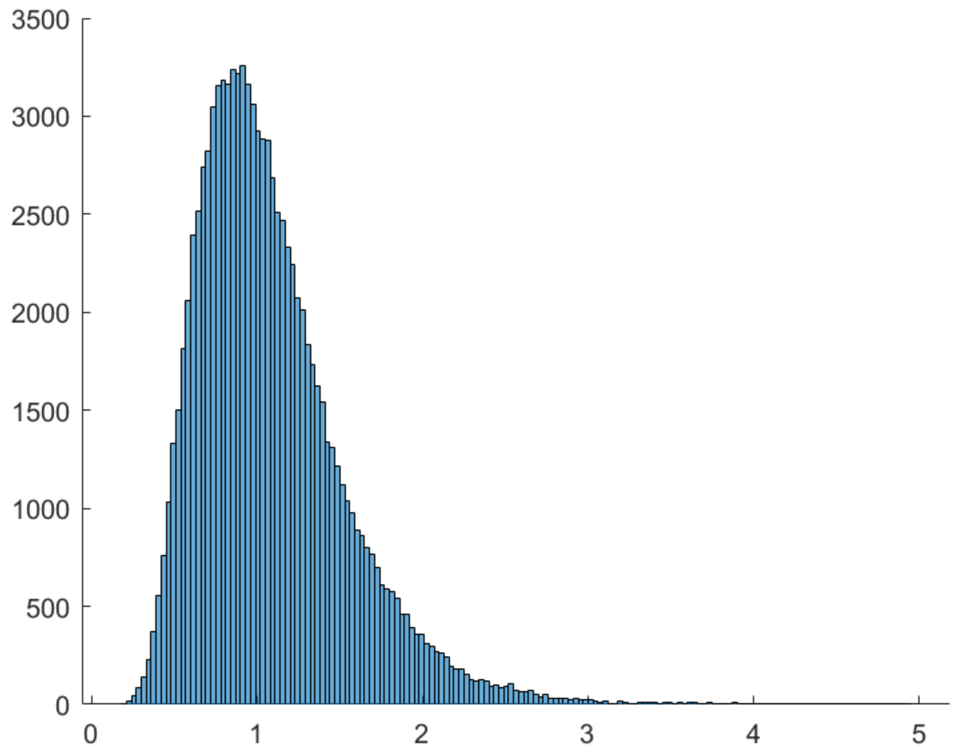
แต่ถ้าเรามาความเร็วเป็นค่าที่โดนยกกำลังดังต่อไปนี้

$$V = e^W$$

เราสามารถให้ค่าของเลขยกกำลัง W มีการแจกแจงแบบปรกติ ผลลัพทธ์ที่ได้คือความเร็วที่มีการกระจายตัวแบบปรกตด้วยล็อก (log-normal distribution) ซึ่งมีฟังก์ชันการแจกแจงของความน่าจะเป็นในรูปแบบดังต่อไปนี้

$$f_V(v) = \frac{1}{v \cdot \sigma_V \sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\frac{\ln(v) - \mu_V}{\sigma_V}\right)^2}$$

```
mu = log(1);
sigma = 0.4;
X = lognrnd(mu,sigma,[1,100000]);
clf;
ax = axes;
hold(ax,'on')
histogram(ax,X)
```


3500
3000
2500
2000
1500
1000
500
0
0
1
2
3
4
5

ความน่าจะเป็นและสัญญาณสุ่ม (Probability & Random Signal)

ตอนที่ 4 : ค่าคาดหวัง ค่าแปรปรวน และ ค่าเบี่ยงเบนมาตรฐาน (Expected Value, Variance, & Standard Deviation)

ค่าคาดหวัง (Expected Value)

หนึ่งในคุณลักษณะของการแจกแจงของความน่าจะเป็นคือค่าคาดหวัง **ค่าคาดหวังนี้ถูกใช้ในการประมาณเหตุการณ์ที่น่าจะเป็นไปได้มากที่สุดโดยอ้างอิงจากฟังก์ชันการแจกแจง**

หากตัวแปรสุ่ม X เป็นแบบไม่ต่อเนื่อง เราสามารถคำนวนหาค่าคาดหวัง $\mathrm{E}\{X\}$ หรือ μ_X ได้ดังต่อไปนี้

$$\mathrm{E}\{X\} = \mu_X = \sum_{\text{all } x} x \cdot f_X(x)$$

ยกตัวอย่างเช่น ค่าคาดหวังของผลรวมของหน้าของลูกเต๋า 2 อันที่ถูกทอยสามารถคำนวณได้ดังต่อไปนี้

$$\begin{aligned}
\mu_X = \mathrm{E}\{X\} &= \sum_{x=2}^{12} x \cdot \left(\frac{6-|x-7|}{36}\right) \\
&= \sum_{x=2}^{7} x \cdot \left(\frac{6-(7-x)}{36}\right) + \sum_{x=8}^{12} x \cdot \left(\frac{6-(x-7)}{36}\right) \\
&= \sum_{x=2}^{7} x \cdot \left(\frac{x-1}{36}\right) + \sum_{x=8}^{12} x \cdot \left(\frac{13-x}{36}\right) = 7
\end{aligned}$$

```
x2_7 = 2:7;
x8_12 = 8:12;
mu_X = sum(x2_7.*(x2_7-1)/36)+sum(x8_12.*(13-x8_12)/36);
```

หากตัวแปรสุ่ม X เป็นแบบต่อเนื่อง เราสามารถคำนวนหาค่าคาดหวัง $\mathrm{E}\{X\}$ หรือ μ_X ได้ดังต่อไปนี้ในรูปแบบของสมการปริพันธ์ดังต่อไปนี้

$$\mu_X = E\{X\} = \int_{-\infty}^{\infty} \{x \cdot f_X(x)\} \mathrm{d}\,x$$

จากนิยามของค่าคาดหวัง คุณสมบัติของค่าคาดหวังมีดังต่อไปนี้

กำหนดให้ k เป็นค่าคงที่ ดังนั้นฟังก์ชันการแจกแจงของค่าคงที่จะอยู่ในรูปของฟังก์ชันแรงดล (impulse function) ดังนี้

$$f_X(x) = \delta(x-k)$$

ดังนั้นเราสามารถหาค่าคาดหวังของค่าคงที่ได้ดังนี้

$$\mathrm{E}\{k\} = \int_{-\infty}^{\infty} \{x \cdot \delta(x-k)\} \mathrm{d}\,x = \int_{-\infty}^{\infty} \{(x+k) \cdot \delta(x)\} \mathrm{d}\,x = k$$

สาเหตุที่เราสามารถหาปริพันธ์เป็นเพราะคุณสมบัติของฟังก์ชันแรงดลดังต่อไปนี้

$$\int_{-\infty}^{\infty} \{f(x) \cdot \delta(x)\} \mathrm{d}\,x = f(0)$$

นอกจากนี้เราสามารถคำนวรหาค่าคาดหวังของตัวแปรสุ่มที่ถูกคูณด้วยค่าคงที่ $Y = k \cdot X$

$$
\begin{aligned}
&\mathrm{E}\{Y\} = \mathrm{E}\{k \cdot X\} \\
&= \int_{-\infty}^{\infty} \{(k \cdot x) \cdot f_X(x)\} \mathrm{d}\, x = k \cdot \int_{-\infty}^{\infty} \{x \cdot f_X(x)\} \mathrm{d}\, x \\
&= k \cdot \mathrm{E}\{X\}
\end{aligned}
$$

หลักการเดียวกันนี้สามารถใช้ในการพิสูจน์หาค่าคาดหวังของผลบวกของหลายฟังก์ชันได้ดังนี้

$$Y = g(X) + h(X)$$

$$\mathrm{E}\{Y\} = \mathrm{E}\{g(X) + h(X)\} = \int_{-\infty}^{\infty} \{(g(x) + h(x)) \cdot f_X(x)\} \mathrm{d}\, x$$

$$= \int_{-\infty}^{\infty} \{g(x) \cdot f_X(x)\} \mathrm{d}\, x + \int_{-\infty}^{\infty} \{h(x) \cdot f_X(x)\} \mathrm{d}\, x$$

$$= \mathrm{E}\{g(X)\} + \mathrm{E}\{h(X)\}$$

ดังนั้นเราสามารถสรุปเป็นสมบัติได้ดังนี้

$$
\begin{bmatrix}
\text{ตัวแปรสุ่ม} & \text{ค่าคาดหวัง} \\
k & k \\
k \cdot X & k \cdot \mathrm{E}\{X\} \\
g(X) + h(X) & \mathrm{E}\{g(X)\} + \mathrm{E}\{h(X)\}
\end{bmatrix}
$$

ค่าแปรปรวน (Variance)

อีกคุณลักษณะหนึ่งของการแจกแจงของความน่าจะเป็นคือค่าแปรปรวนซึ่งบอกถึงการกระจายตัวของค่าที่ตะวแปรสามารถเป็นได้ **ยิ่งค่าแปรปรวนมาก การแจกแจงความน่าจะเป็นก็จะกว้างตามไปด้วย ที่สำคัญที่สุดคือค่าแปรปรวนนั้นจะต้องมากกว่าหรือเท่ากับศูนย์เสมอ**

หากกำหนดให้ X เป็นตัวแปรสุ่ม เราสามารถคำนวณหาความแปรปรวน $\mathrm{Var}\{X\}$ หรือ σ_X^2 ได้กังต่อไปนี้

$$
\begin{aligned}
&\sigma_X^2 = \mathrm{Var}\{X\} = \mathrm{E}\left\{(X - \mu_X)^2\right\} \\
&= \mathrm{E}\left\{X^2 - 2X\mu_X + \mu_X^2\right\} \\
&= \mathrm{E}\{X^2\} - 2\mu_X \cdot E\{X\} + \mathrm{E}\left\{\mu_X^2\right\} \\
&= \mathrm{E}\{X^2\} - \mu_X^2
\end{aligned}
$$

ยกตัวอย่างเช่น ค่าแปรปรวนของผลรวมของหน้าของลูกเต๋า 2 อันที่ถูกทอยสามารถคำนวณได้ดังต่อไปนี้

$$
\begin{aligned}
&\sigma_X^2 = \mathrm{Var}\{X\} = \mathrm{E}\{X^2\} - \mu_X^2 \\
&= \sum_{x=2}^{12} x^2 \cdot \left(\frac{6 - |x - 7|}{36}\right) - 7^2 \\
&= \sum_{x=2}^{7} x^2 \cdot \left(\frac{6 - (7 - x)}{36}\right) + \sum_{x=8}^{12} x^2 \cdot \left(\frac{6 - (x - 7)}{36}\right) - 49 \\
&= \sum_{x=2}^{7} x^2 \cdot \left(\frac{x - 1}{36}\right) + \sum_{x=8}^{12} x^2 \cdot \left(\frac{13 - x}{36}\right) - 49 = \frac{35}{6}
\end{aligned}
$$

```
x2_7 = 2:7;
x8_12 = 8:12;
```

```
mu_X = sum(x2_7.^2.*(x2_7-1)/36)+sum(x8_12.^2.*(13-x8_12)/36)-49;
```

จากนิยามของค่าแปรปรวนที่กล่าวมา เราสามารถหาสมบัติของค่าแปรปรวนได้ดังต่อไปนี้

กำหนดให้ค่า k เป็นค่าคงที่ ค่าแปรปรวนนั้นจะเท่ากับศูนย์เนื่องจากค่าคงที่เป็นค่าที่ทราบอยู่แล้วและไม่แปรปรวน ซึ่งสามารถพิสูจน์ได้ดังนี้

$$\sigma_X^2 = \mathrm{Var}\{k\} = \mathrm{E}\{k^2\} - k^2 = k^2 - k^2 = 0$$

อีกสมบัติหนึ่งที่เราสามารถพิสูจน์จากนิยามของค่าคาดหวังและค่าแปรปรวนคือค่าแปรปรวนของตัวแปรสุ่มที่ถูกคูณด้วยค่าคงที่ ซึ่งสามารถพิสูจน์ได้ดังต่อไปนี้

$$\begin{aligned}
&\sigma_X^2 = \mathrm{Var}\{k \cdot X\} = \mathrm{E}\{(k \cdot X)^2\} - (k \cdot \mu_X)^2 \\
&= \mathrm{E}\{k^2 \cdot X^2\} - k^2 \cdot \mu_X^2 = k^2 \cdot \mathrm{E}\{X^2\} - k^2 \cdot \mu_X^2 \\
&= k^2 \cdot \left(\mathrm{E}\{X^2\} - \mu_X^2\right) = k^2 \cdot \mathrm{Var}\{X\}
\end{aligned}$$

ความน่าจะเป็นและสัญญาณสุ่ม (Probability & Random Signal)

ตอนที่ 5 : ตัวแปรสุ่มหลายตัวแปร (Multiple Random Variables)

ความน่าจะเป็นร่วม (Joint Probability)

เราใช้ตัวแปรสุ่มหนึ่งตัวในการอธิบายเหตุการณ์หนึ่งเหตุการณ์ในกรณีต่างๆ แต่**เราสามารถอธิบายเหตุการณ์หลายๆเห็นการที่เกิดขึ้นพร้อมกันโดยใช้ตัวแปรสุ่มมากกว่าหนึ่งตัว** เช่น เหตุการณ์ที่ฟ้าครึ้มและฝนตกพร้อมกัน หรือ เหตุการณ์ที่ไพ่ใบแรกเป็น King และไพ่ใบที่สองเป็น Queen **เราสามารถใช้ตรรกศาสตร์ในการอธิบายเหตุการณ์ที่เกิดขึ้นพร้อมกัน หากกำหนดให้เหตุการณ์แรกเป็น E_1 และเหตุการณ์ที่สองเป็น E_2** เราจะเขียนเหตุการณ์และความน่าจะเป็นที่ทั้งสองเหตุการณ์เกิดขึ้นพร้อมกันได้ดังต่อไปนี้

$$E = E_1 \wedge E_2$$
$$\mathrm{P}(E) = \mathrm{P}(E_1 \wedge E_2)$$

เราเรียกความน่าจะเป็นนี้ว่าความน่าจะเป็นร่วม (joint probability) ซึ่งความสัมพันธ์ของความน่าจะเป็นสามารถสรุปเป็นตารางได้ดังต่อไปนี้

เหตุการณ์	E_1	$\neg E_1$	$E_1 \vee \neg E_1 = \text{True}$
E_2	$\mathrm{P}(E_1 \wedge E_2)$	$\mathrm{P}(\neg E_1 \wedge E_2)$	$\mathrm{P}(E_2) = \mathrm{P}(E_1 \wedge E_2) + \mathrm{P}(E_1 \wedge \neg E_2)$
$\neg E_2$	$\mathrm{P}(E_1 \wedge \neg E_2)$	$\mathrm{P}(\neg E_1 \wedge \neg E_2)$	$\mathrm{P}(\neg E_2) = P(\neg E_1 \wedge E_2) + \mathrm{P}(\neg E_1 \wedge \neg E_2)$
$E_2 \vee \neg E_2 = \text{True}$	$\mathrm{P}(E_1) = P(E_1 \wedge E_2) + \mathrm{P}(E_1 \wedge \neg E_2)$	$\mathrm{P}(\neg E_1) = P(\neg E_1 \wedge E_2) + \mathrm{P}(\neg E_1 \wedge \neg E_2)$	$\mathrm{P}(E_1) + \mathrm{P}(\neg E_1) = \mathrm{P}(E_2) + \mathrm{P}(\neg E_2) = 1$

ความน่าจะเป็นแบบมีเงื่อนไข (Conditional Probability)

ในหลายครั้ง **เราสามารถที่จะสังเกตเหตุการณ์ (observe) เพื่อใช้ในการคาดเดาว่าอีกเหตุการณ์หนึ่งจะเกิดขึ้นหรือไม่** เช่น ความน่าจะเป็นที่มีคนมาเยี่ยมบ้านเมื่อเราได้ยินเสียงสุนัขที่บ้านเห่า หรือ ความน่าจะเป็นที่อาชญากรรมจะเกิดขึ้น โดยที่เราเห็นโคนันกับแก๊งค์นักสืบเยาวชนในบริเวณนั้น

ความน่าจะเป็นแบบดังกล่าวถูกเรียกว่า **ความน่าจะเป็นแบบมีเงื่อนไข** (conditional probability)

ความน่าจะเป็นแบบเงื่อนไขนี้สามารถถูกคำนวณได้จากความสัมพันธ์ดังต่อไปนี้

$$\mathrm{P}(E_1 | E_2) = \frac{\mathrm{P}(E_1 \wedge E_2)}{\mathrm{P}(E_2)}$$

โดยที่

$\mathrm{P}(E_1|E_2)$ คือความน่าจะเป็นที่เหตุการณ์ E_1 เกิดขึ้นโดยกำหนดให้ เหตุการณ์ E_2 เกิดขึ้น

$\mathrm{P}(E_1 \wedge E_2)$ คือความน่าจะเป็นร่วมเหตุการณ์ E_1 เกิดขึ้นพร้อมกับ เหตุการณ์ E_2

จากตารางและนิยามของความน่าเป็นแบบเงื่อนไข เราสามารถหาความสัมพันธ์ของความน่าจะเป็นดังกล่าวได้

$$\mathrm{P}(E_1) = \mathrm{P}(E_1 \wedge E_2) + \mathrm{P}(E_1 \wedge \neg E_2) = \mathrm{P}(E_1|E_2) \cdot \mathrm{P}(E_2) + \mathrm{P}(E_1|\neg E_2) \cdot \mathrm{P}(\neg E_2)$$

นอกจากนี้ **เราสามารถหาความสัมพันธ์ของความน่าจะเป็นแบบเงื่อนไข ได้ในรูปของทฤษฎีของ (Bayes' theorem)** ได้ดังต่อไปนี้

$$\mathrm{P}(E_1|E_2) = \frac{\mathrm{P}(E_2|E_1) \cdot \mathrm{P}(E_1)}{\mathrm{P}(E_2)} = \frac{\mathrm{P}(E_2|E_1) \cdot \mathrm{P}(E_1)}{\mathrm{P}(E_2|E_1) \cdot \mathrm{P}(E_1) + \mathrm{P}(E_2|\neg E_1) \cdot \mathrm{P}(\neg E_1)}$$

ตัวอย่างที่ 1 : Am I being robbed ?

ณ ธนาคารเล็กๆแห่งหนึ่ง นายธนาคารได้จ้างรปภ.คนใหม่ที่มีประสิทธิภาพสูง ถ้าไม่เกิดเหตุการณ์ปล้นในธนาคารภายในหนึ่งกะ นายรปภ.คนดังกล่าวจะมีโอกาสหลับที่ 1% ($S = 0$ คือรปภ. ไม่หลับ $S = 1$ คือรปภ. หลับ) จะอย่างไรก็ตาม นายธนาคารเป็นคนรอบคอบ เขาได้วิเคราะห์ข้อมูลจากธนาคารสาขาอื่นๆ ทำให้เขาทราบว่าความน่าจะเป็นที่ธนาคารจะโดนปล้นต่อหนึ่งกะคือ 10% ($R = 0$ คือไม่โดนปล้น $R = 1$ คือโดนปล้น)

ในรถตู้ที่อยู่บนถนนถัดไปมีกลุ่มโจรกำลังวางแผนโจรกรรมกับธนาคารแห่งนี้ โจรกลุ่มนี้ตัดสินใจที่จะใช้ลูกดอกยาสลบซึ่งมีผลทำให้คนหลับได้ทันทีแต่มือปืนของฝ่ายโจรบอกกับคนวางแผนว่าเขาอาจจะเผลอยิงโดนหัวเข็มขัดหรือสิ่งของอื่นๆได้ซึ่งโอกาสในการยิงโดนรปภ.จริงๆคือ 95% กล่าวอีกอย่างคือ มีโอกาส 95% ที่รปภ.จะหลับถ้าเกิดการปล้น

เมื่อนายธนาคารถึงเวลาออกมาตรวจเช็คงาน นายธนาคารเห็นสภาพรปภ.กำลังนั่งหลับอยู่ คำถามคือความน่าจะเป็นที่ธนาคารโดนปล้นอยู่เป็นเท่าไหร่

คำถามที่ถามมานั้นถามถึงความน่าจะเป็นแบบเงื่อนไขซึ่งสามารถเขียนได้ดังต่อไปนี้

$$\mathrm{P}(R = 1 | S = 1) : \text{ความน่าจะเป็นที่ธนาคารกำลังโดนปล้นเมื่อเห็นว่ารปภ. หลับอยู่}$$

เราสามารถใช้ความสัมพันธ์และทฤษฎีของเบย์ในการคำนวณหาความน่าจะเป็นที่ธนาคารกำลังโดนปล้นเมื่อเห็นว่ารปภ. หลับอยู่ ได้ดังนี้

$$\begin{aligned}
\mathrm{P}(R = 1 | S = 1) &= \frac{\mathrm{P}(R = 1 \wedge S = 1)}{\mathrm{P}(S = 1)} = \frac{\mathrm{P}(S = 1 | R = 1) \cdot \mathrm{P}(R = 1)}{\mathrm{P}(S = 1)} \\
&= \frac{\mathrm{P}(S = 1 | R = 1) \cdot \mathrm{P}(R = 1)}{\mathrm{P}(S = 1 | R = 1) \cdot \mathrm{P}(R = 1) + \mathrm{P}(S = 1 | R = 0) \cdot (1 - \mathrm{P}(R = 1))} \\
&= \frac{0.95 \cdot 0.1}{0.95 \cdot 0.1 + 0.01 \cdot (1 - 0.1)} \approx 0.9135
\end{aligned}$$

```
S1R0 = 0.01;
R1 = 0.1;
S1R1 = 0.95;

R1S1 = S1R1*R1/(S1R1*R1+S1R0*(1-R1))
```

```
R1S1 = 0.9135
```

นั่นหมายความว่าการที่เห็นว่ารปภ.หลับอยู่อาจจะมีโอกาส 91.35% ธนาคารกำลังโดนปล้นอยู่

ตัวแปนที่อิสระต่อกัน (Indepedent Random Variables)

หากเหตุการณ์ทั้งสองเหตุการณ์นั้นเป็นอิสระต่อกัน (independent) เหตุการณ์หนึ่งจะไม่ส่งผลอะไรให้อีกเหตุการณ์หนึ่งเลย ซึ่งสามารถเขียนเป็นสมการความสัมพันธ์ได้ดังต่อไปนี้

$$\mathrm{P}(E_1 | E_2) = \mathrm{P}(E_1)$$

จากสมการดังกล่าวและทฤษฎีของเบย์ **เราสามารถที่จะคำนวณหาความน่าจะเป็นร่วมของทั้งสองเหตุการณ์ที่เป็นอิสระต่อกันได้ดังนี้**

$$\mathrm{P}(E_1 \wedge E_2) = \mathrm{P}(E_1) \cdot \mathrm{P}(E_2)$$

ตัวอย่างที่ 2 : โยนเหรียญ

กำหนดให้เหรียญที่ถูกโยนมีความน่าจะเป็นที่ออกหัวหรือก้อยเท่าๆกัน หากเราโยนเหรียญ 10 ครั้งติดต่อกัน ความน่าจะเป็นจะออกหัวอย่างน้อย 1 ครั้งเป็นเท่าไหร่

กำหนดให้ X_k เป็นตัวแปรที่บอกถึงผลของการโยนเหรียญครั้งที่ k($X_k = 0$ เหรียญที่โยนออกก้อย $X_k = 1$ เหรียญที่โยนออกกหัว)

เราสามารถหาความน่าจะเป็นของแต่ละการโยนออกเป็นก้อยได้ดังต่อไปนี้

$$\mathrm{P}(X_k = 0) = 0.5$$

ความน่าจะเป็นที่ทุกการโยนมีผลออกมาเป็นก้อย $E_{\mathrm{AT}} = (X_1 = 0) \wedge (X_2 = 0) \wedge \cdots \wedge (X_{10} = 0)$ สามารถเขียนอยู่ในรูปของผลคูณความน่าจะเป็นในการโยนของแต่ละครั้งเนื่องจากว่าการโยนในแต่ละครั้งเป็นอิสระต่อกัน ดังนั้น

$$\mathrm{P}(E_{\mathrm{AT}}) = \mathrm{P}(X_1 = 0) \cdot \mathrm{P}(X_2 = 0) \cdot \cdots \cdot \mathrm{P}(X_{10} = 0) = (0.5)^{10}$$

กำหนดให้ E_H เป็นเหตุการณ์ที่โยนเหรียญทั้งหมด 10 เหรียญและผลออกมาเป็นหัวอย่างน้อย 1 ครั้งนั้นซึ่งเป็นเหตุการณ์ที่ตรงกันข้ามกับเหตุการณ์ที่โยนเหรียญทั้งหมด 10 เหรียญและผลออกมาเป็นก้อยทั้ง 10 ครั้ง

$$\mathrm{P}(E_H) = 1 - \mathrm{P}(E_{\mathrm{AT}}) = 1 - (0.5)^{10} \approx 0.9990$$

ฟังก์ชันการแจกแจงของความน่าจะเป็นร่วม (joint probability distribution function)

เช่นเดียวกับตัวแปรเดี่ยว **เราสามารถใช้ฟังก์ชันในการอธิบายการแจกแจงของความน่าจะเป็นร่วมโดยใช้ฟังก์ชันหลายตัวแปร** (multi-variable function)

หากเรากำหนดให้ X และ Yเป็นตัวแปรสุ่มแบบไม่ต่อเนื่อง เราสามารถเขียนความสัมพันธ์ของความน่าจะเป็นได้ดังต่อไปนี้

$$\mathrm{P}((X = x) \wedge (Y = y)) = f_{XY}(x, y)$$

ซึ่งฟังก์ชันการแจกแจงของตัวแปรสุ่มแบบไม่ต่อเนื่องต้องมีพฤติกรรมดังต่อไปนี้

$$\begin{aligned}&\sum_{\text{for all } y} \sum_{\text{for all } x} f_{XY}(x, y) f_{XY}(x, y) = 1 \\ &0 \leq f_{XY}(x, y) \leq 1\end{aligned}$$

หากเรากำหนดให้ X และ Yเป็นตัวแปรสุ่มแบบต่อเนื่อง เราสามารถเขียนความสัมพันธ์ของความน่าจะเป็นได้ดังต่อไปนี้

$$\mathrm{P}((X \leq k) \wedge (Y \leq h)) = \int_{-\infty}^{h} \int_{-\infty}^{k} f_{XY}(x, y)\, \mathrm{d}\, x\, \mathrm{d}\, y$$

ี่ึ่งฟังก์ชันการแจกแจงของตัวแปรสุ่มแบบไม่ต่อเนื่องต้องมีพฤติกรรมดังต่อไปนี้

$$\begin{aligned}&\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y)\, \mathrm{d}\, x\, \mathrm{d}\, y = 1 \\ &f_{XY}(x, y) \geq 0\end{aligned}$$

หากตัวแปรทั้งสองตัวเป็นอิสระต่อกัน ฟังก์ชันการแจกแจงสามารถถูกเขียนในรูปของผลคูณได้ดังต่อไปนี้

$$f_{XY}(x, y) = f_{XY}(x) \cdot f_{XY}(y)$$

ความน่าจะเป็นและสัญญาณสุ่ม (Probability & Random Signal)

ตอนที่ 6 : ความสัมพันธ์ระหว่างตัวแปรสุ่มหลายตัว (Relationship between Multiple Random Variables)

ค่าแปรปรวนร่วมเกี่ยว (Covariance)

ค่าของตัวแปรสุ่มสองตัวสามารถแปรปรวนร่วมกันได้ เราเรียกค่านี้ว่าค่าแปรปรวนร่วมเกี่ยว (covariance)

หากเรากำหนดให้ X และ Y เป็นตัวแปรสุ่ม ค่าแปรปรวนร่วมเกี่ยว $\mathrm{Cov}\{X,Y\}$ หรือ σ_{XY} ของตัวแปร X และ Y จะถูกนิยามดังต่อไปนี้

$$\begin{aligned}
\sigma_{XY} &= \mathrm{Cov}\{X,Y\} = \mathrm{E}\{(X-\mu_X)\cdot(Y-\mu_Y)\} \\
&= \mathrm{E}\{X\cdot Y-\mu_X\cdot Y-X\cdot\mu_Y+\mu_X\cdot\mu_Y\} \\
&= \mathrm{E}\{X\cdot Y\}-\mu_X\cdot\mathrm{E}\{Y\}-\mu_Y\cdot E\{X\}+\mu_X\cdot\mu_Y \\
&= \mathrm{E}\{X\cdot Y\}-\mu_X\cdot\mu_Y
\end{aligned}$$

ซึ่งความแปรปรวนร่วมเกี่ยวจะมีสมบัติดังต่อไปนี้

$$\begin{aligned}
&\mathrm{Cov}\{X,Y\} = \mathrm{Cov}\{Y,X\} \\
&|\mathrm{Cov}\{X,Y\}| \le \sqrt{\mathrm{Var}\{X\}\cdot\mathrm{Var}\{Y\}}
\end{aligned}$$

ค่าแปรปรวนร่วมเกี่ยวนั้นอธิบายถึงความแปรปรวนที่สองตัวแปรนั้นแปรเปลี่ยนไปในทางเดียวกัน

ค่าสหสัมพันธ์ (Correlation)

อีกหนึ่งค่าที่สามารถอธิบายความสัมพันธ์ระหว่างตัวแปรสุ่มสองตัวได้คือค่าสหสัมพันธ์ (correlation)

หากเรากำหนดให้ X และ Y เป็นตัวแปรสุ่ม ค่าสหสัมพันธ์ของตัวแปรสุ่มทั้งสอง ρ_{XY} หรือ $\mathrm{Cor}\{X,Y\}$ จะถูกนิยามดังต่อไปนี้

$$\begin{aligned}
&\rho_{XY} = \mathrm{Cor}\{X,Y\} = \frac{\sigma_{XY}}{\sigma_X\cdot\sigma_Y} \\
&-1 \le \mathrm{Cor}\{X,Y\} \le 1
\end{aligned}$$

หากเรามีชุดข้อมูลเป็นคู่อันดับ $\langle x, y\rangle$ เราสามารถใช้ฟังก์ชัน corrcoef เพื่อคำนวณหาค่าสหสัมพันธ์ระหว่างข้อมูลชุด x และข้อมูลชุด y ซึ่งจะได้ผลลัพธ์เป็นเมตริกซ์ดังต่อไปนี้

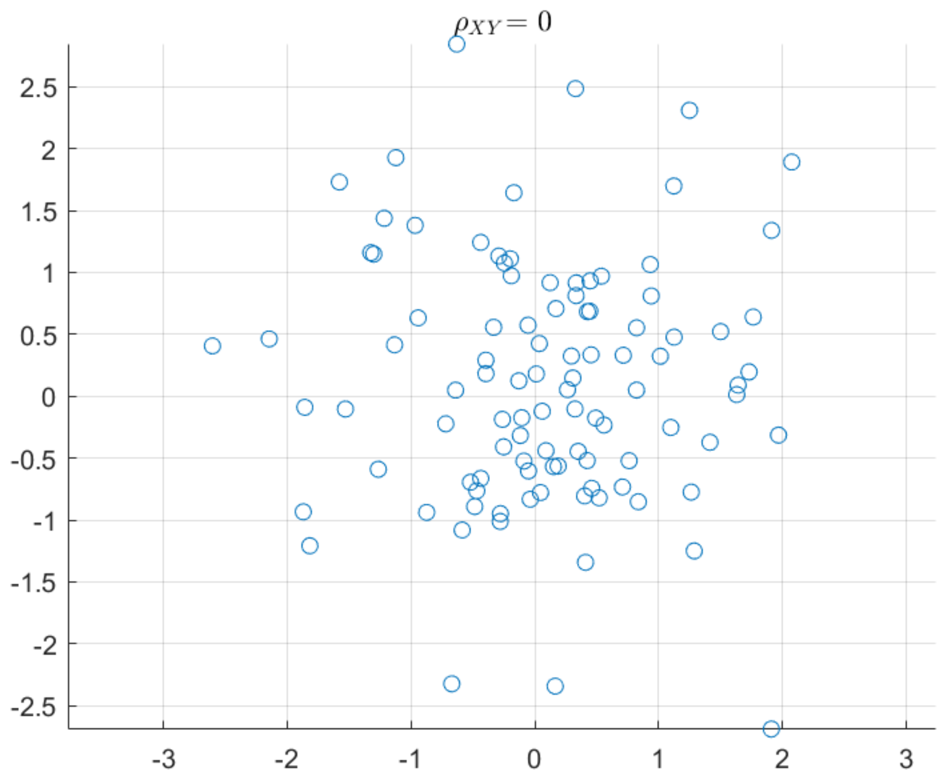
$$\mathrm{corrcoef}(X,Y) = \begin{bmatrix} \mathrm{Cor}\{X,X\} & \mathrm{Cor}\{X,Y\} \\ \mathrm{Cor}\{X,Y\} & \mathrm{Cor}\{Y,Y\} \end{bmatrix}$$

ซึ่ง $\mathrm{Cor}\{X,X\} = \mathrm{Cor}\{Y,Y\} = 1$

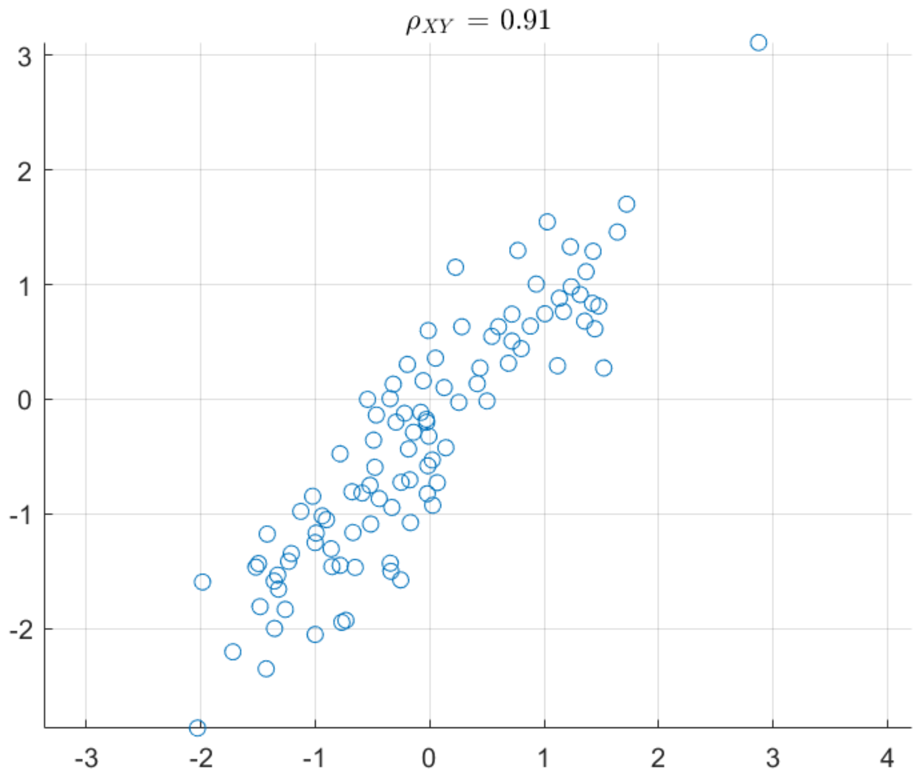
เราจะใช้สมาชิกตัวที่ไม่ได้อยู่ในแนวทแยงเท่านั้นเนื่องจากค่าสหพันธ์ของค่าเดียวกันจะมีผลเป็น 0 เสมอ

```
x = randn(1,100);
y = randn(1,100);
clf
ax = axes;
hold(ax,'on')
grid(ax,'on')
plot(ax,x,y,'o')
axis(ax,'equal')
```

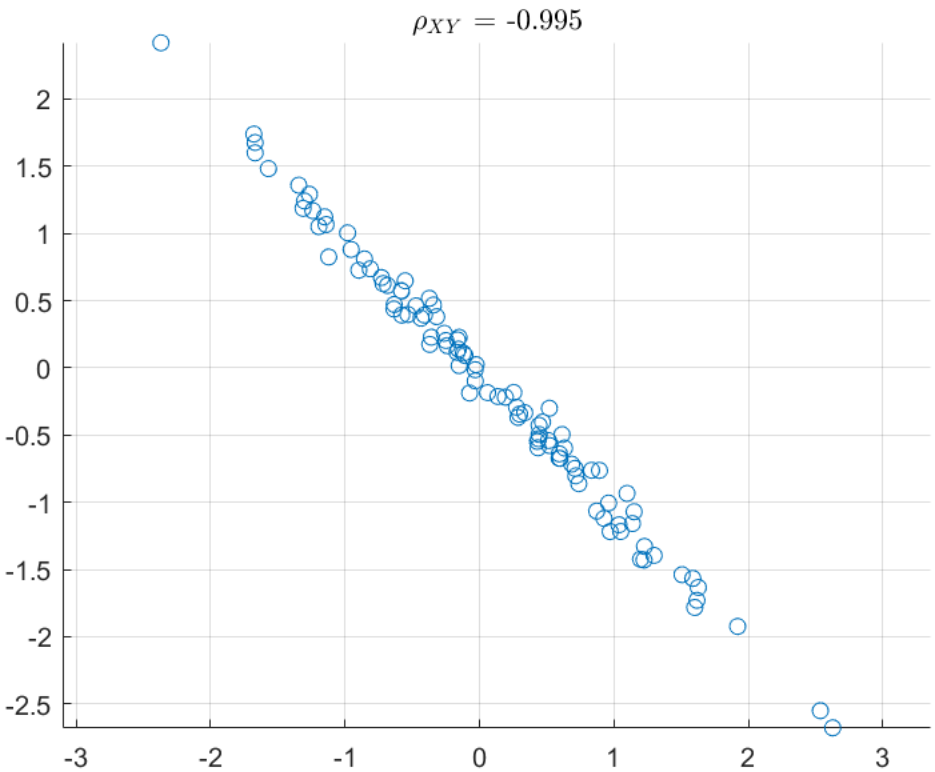
```
title(ax,'$\rho_{XY}$= 0','Interpreter',"latex")
```




```
x = randn(1,100);
y = x+0.5*(randn(size(x))-0.5);
clf
ax = axes;
hold(ax,'on')
grid(ax,'on')
plot(ax,x,y,'o')
axis(ax,'equal')
C = corrcoef(x,y);
title(ax,sprintf('$\\rho_{XY}$ = %.3f',C(1,2)),'Interpreter',"latex")
```




```
x = randn(1,100);
y = -x+0.1*(randn(size(x))-0.5);
clf
ax = axes;
hold(ax,'on')
grid(ax,'on')
plot(ax,x,y,'o')
axis(ax,'equal')
C = corrcoef(x,y);
title(ax,sprintf('$\\rho_{XY}$ = %.3f',C(1,2)),'Interpreter',"latex")
```



สมบัติของค่าคาดหวังและค่าแปรปรวน (Properties of Expected Value & Variance)

ในกรณีที่เรามีค่าที่ไม่แน่นอนมากกว่าหนึ่งค่ามาผสมรวมกัน เราจะได้ค่าที่มีความไม่แน่นอนใหม่ ทั้งค่าคาดหวังและค่าแปรปรวนก็จะเปลี่ยนตามด้วย

หากกำหนดให้ X, Y, และ Z เป็นตัวแปรสุ่ม และ Zเป็นผลรวมเชิงเส้นของตัวแปร X และ Y ดังต่อไปนี้

$$Z = \alpha \cdot X + \beta \cdot Y$$

เราสามารถคำนวณหาค่าคาดหวังของตัวแปร Z ดังต่อไปนี้

$$\mu_Z = E\{Z\} = \mathrm{E}\{\alpha \cdot X + \beta \cdot Y\} = \alpha \cdot \mu_X + \beta \cdot \mu_Y$$

นอกจากนี้ เราสามารถคำนวณหาค่าแปรปรวนของตัวแปร Z ดังต่อไปนี้

$$\sigma_Z^2 = \mathrm{Var}\{Z\} = \alpha^2 \cdot \sigma_X^2 + 2\alpha\beta \cdot \sigma_{XY} + \beta^2 \cdot \sigma_Y^2$$
$$\sigma_Z^2 = \mathrm{Var}\{Z\} = \alpha^2 \cdot \sigma_X^2 + 2\alpha\beta \cdot \rho_{XY} \cdot \sigma_X \cdot \sigma_Y + \beta^2 \cdot \sigma_Y^2$$

ความน่าจะเป็นและสัญญาณสุ่ม (Probability & Random Signal)

ตอนที่ 7 : เวกเตอร์สุ่ม (Random Vector)

ในกรณีที่เรามีตัวแปรสุ่มมากกว่าหนึ่งตัว **เราสามารถจัดเรียงตัวแปรเหล่านี้ให้อยู่ในรูปของเวกเตอร์** เราจะกำหนดให้ ตัวแปรสุ่มมีทั้งหมด n ตัว และกำหนดให้เวกเตอร์สุ่ม (random vector) $\mathbf{X}$ ประกอบไปด้วยสมาชิกดังนี้

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$$

หากตัวแปรสุ่มแต่ละตัวเป็นแบบไม่ต่อเนื่อง เราสามารถจำลองความไม่แน่นอนของเวกเตอร์ดังกล่าวได้ว่าการแจกแจงซึ่งมีหลากหลายรูปแบบ ซึ่งล้วนแต่ต้องเขียนอยู่ในรูปของฟังก์ชันต่อไปนี้

$$\mathrm{P}(\mathbf{X} \in A) = \sum_{\mathbf{x} \in A} \cdots \sum f_X(\mathbf{x})$$

โดยที่

$$\sum_{\text{for all } \mathbf{x}} \cdots \sum f_{\mathbf{X}}(\mathbf{x}) = 1$$
$$0 \le f_{\mathbf{X}}(\mathbf{x}) \le 1$$

หากตัวแปรสุ่มแต่ละตัวเป็นแบบต่อเนื่อง เราสามารถเขียนฟังก์ชันแจกแจงได้ในรูปแบบดังต่อไปนี้

$$\mathrm{P}(\mathbf{X} \in A) = \int^{A} \left\{ \cdots \int \{f_X(x_1, \cdots, x_n)\}\, \mathrm{d}\, x_1 \cdots \right\} \mathrm{d}\, x_n$$

โดยที่

$$\int_{-\infty}^{\infty} \left\{ \cdots \int_{-\infty}^{\infty} \{f_X(x_1, \cdots, x_n)\}\, \mathrm{d}\, x_1 \cdots \right\} \mathrm{d}\, x_n = 1$$
$$0 \le f_{\mathbf{X}}(\mathbf{x}) \le 1$$

ถึงแม้เวกเตอร์สุ่มเป็นเวกเตอร์หลัก **(column) ในการเขียนสมการคณิตศาสตร์ แต่ในการเขียนโปรแกรมส่วนใหญ่ เรามักจะเรียกเวกเตอร์ให้เป็นแนวนอนเพื่อที่จะให้ค่าของเวกเตอร์นั้นเรียงกันได้ในแนวตั้ง** เช่น กำหนดให้ตัวแปรสุ่มมีทั้งหมด 3 ตัวซึ่งจะถูกเก็บค่า/สุ่มค่าทั้งหมด 4 ครั้ง เราจะได้ชุดข้อมูลในรูปแบบดังกล่าว

$$\text{ข้อมูล : } \begin{bmatrix} \mathbf{X}_{(1)} & \mathbf{X}_{(2)} & \mathbf{X}_{(3)} & \mathbf{X}_{(4)} \end{bmatrix}$$

ซึ่งสามารถเขียนเป็นโปรแกรมได้ดังนี้

```
X = rand(4,3) % uniform distribution between 0 & 1
```

```
X = 4×3
    0.4844    0.8296    0.2297
    0.0274    0.5490    0.5741
    0.9214    0.1532    0.9114
    0.5770    0.2602    0.5215
```

ค่าคาดหวังและเมตริกซ์แปรปรวน

ไม่ว่าการกระจายตัวจะเป็นอย่างไร เราสามารถที่จะจำลองลักษณะความไม่แน่นอนจากคุณลักษณะต่างๆได้ หนึ่งในนั้นก็คือ**ค่าคาดหวังของเวกเตอร์สุ่ม**ซึ่งเท่ากับเวกเตอร์ของค่าคาดหวังของแต่ละสมาชิกในเวกเตอร์

$$\mu_{\mathbf{X}} = E\{\mathbf{X}\} = \begin{bmatrix} E\{X_1\} \\ \vdots \\ E\{X_n\} \end{bmatrix}$$

นอกจากนี้ เรายังสามารถอธิบายความแปรปรวนของเวกเตอร์สุ่มได้ เช่นเดียวกันกับกรณีของตัวแปรสุ่มสองตัว เราจำเป็นที่จะต้องจับคู่ระหว่างตัวแปรสุ่มเพื่อคำนวณหาค่าแปรปรวนร่วมเกี่ยวและยังต้องหาค่าแปรปรวนของแต่ละตัวแปรสุ่มอีกด้วย

หากเรามีจำนวนตัวแปรสุ่มทั้งหมด n ตัวด้วยกัน เราจะมีค่าแปรปรวนทั้งหมด n ตัว $(\sigma_{X_i}^2)$ และค่าแปรปรวนร่วมเกี่ยวเป็นจำนวนทั้งหมด $n \cdot (n-1)$ ตัว $(\sigma_{X_i X_j})$ เพื่อความสะดวกในการเขียนอธิบาย **เราสามารถจัดค่าแปรปรวนทั้งหมดนี้ให้อยู่ในรูปของเมตริกซ์ดังต่อไปนี้**

$$\mathbf{K_X} = \mathrm{Cov}\{\mathbf{X}, \mathbf{X}\} = \begin{bmatrix} \sigma_{X_1}^2 & \sigma_{X_1X_2} & \cdots & \sigma_{X_1X_n} \\ \sigma_{X_1X_2} & \sigma_{X_2}^2 & \cdots & \sigma_{X_2X_n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{X_1X_n} & \sigma_{X_2X_n} & \cdots & \sigma_{X_n}^2 \end{bmatrix}$$

เมตริกซ์นี้ถูกเรียกว่า เมตริกซ์ค่าแปรปรวน-ค่าแปรปรวนร่วมเกี่ยว หรือเรียกสั้นๆว่า **เมติกซ์ค่าแปรปรวนร่วมเกี่ยว** (variance-covariance matrix/covariance matrix) ซึ่งเมตริกซ์นี้จะมีคุณสมบัติดังต่อไปนี้

$\mathbf{v}^\top \cdot \mathbf{K_X} \cdot \mathbf{v} \geq 0$

$\mathbf{K_X} = \mathbf{K_X^\top}$

$\mathrm{Cov}\{\mathbf{A} \cdot \mathbf{X} + \mathbf{b}, \mathbf{A} \cdot \mathbf{X} + \mathbf{b}\} = \mathbf{A} \cdot \mathbf{K_X} \cdot \mathbf{A}^\top$

สองสมบัติแรกแสดงให้เห็นว่าเราไม่สามารถที่จะเลือกจำลองเมตริกซ์ค่าแปรปรวนแบบสุ่มสี่สุ่มห้าได้

หากเรามีชุดข้อมูล เราสามารถที่จะใช้ฟังก์ชัน cov ของ MATLAB ในการหาเมตริกซ์ค่าแปรปรวนนี้ได้เลย

```
X_U = rand(100,2);
X = [X_U 2*X_U+0.1*(rand(size(X_U))-0.5)];

K = cov(X)
```

```
K = 4×4
    0.0835   -0.0130    0.1675   -0.0251
   -0.0130    0.0865   -0.0272    0.1725
    0.1675   -0.0272    0.3370   -0.0530
   -0.0251    0.1725   -0.0530    0.3450
```

```
R = corrcoef(X); % เมตริกซ์สหสัมพันธ์
```

การแจกแจงแบบปรกติหลายตัวแปร (Multivariate Normal Distribution)

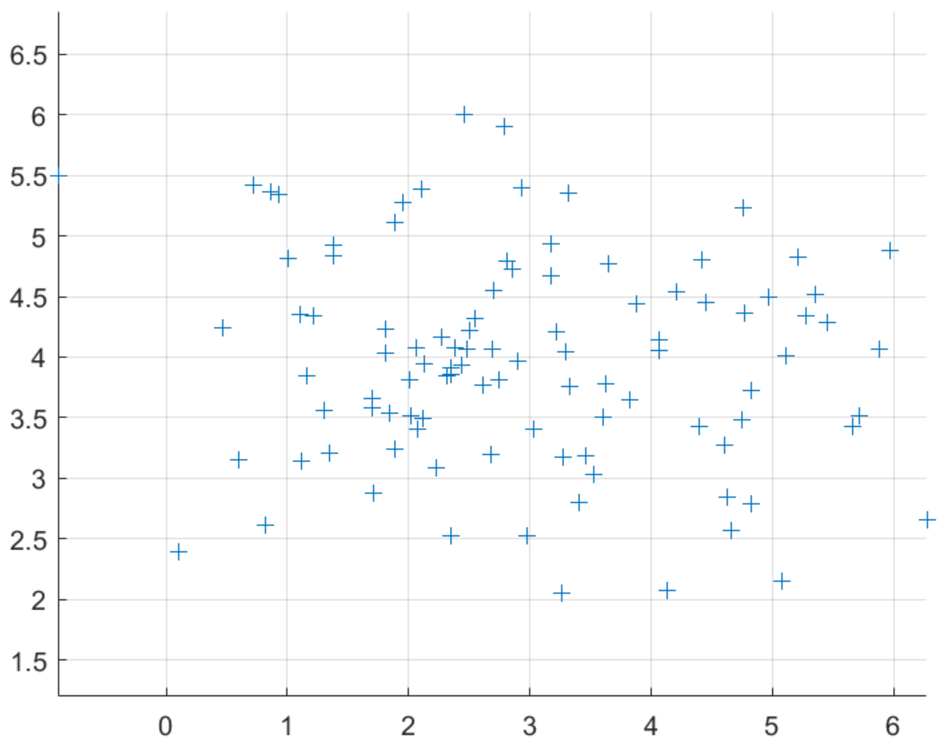
การแจกแจงที่มักจะถูกใช้บ่อยครั้งในการจำลองความไม่แน่นอนที่มีหลายตัวแปรคือการแจกแจงแบบปรกติสำหรับหลายตัวแปร (Multivariate Normal Distribution) ซึ่งสามารถเขียนเป็นฟังก์ชันได้ดังนี้

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{K}_{\mathbf{X}}|}} e^{-\frac{1}{2}(\mathbf{x}-\mu_{\mathbf{X}})^{\top}\mathbf{K}_{\mathbf{X}}^{-1}(\mathbf{x}-\mu_{\mathbf{X}})}$$

การแจกแจงนี้ถูกใช้เป็นแบบจำลองความไม่แน่นอนของพฤติกรรมหลายอย่างในระบบหุ่นยนต์เช่นความไม่แน่นอนของตำแหน่งในพื้นที่ 2 มิติ หรือความไม่แน่นอนของค่าต่างๆที่อ่านได้จากเซนเซอร์

เราสามารถใช้ฟังก์ชัน mvnrnd ของ MATLAB ในการสุ่มค่าของตัวแปรดังต่อไปนี้

```
x_est = [3;4];
var = [2;1];
covar = 2*sqrt(abs(prod(var)))*(rand-1/2);
P = diag(var)+[0 covar; covar 0]; % must be semi-positive definite
X = mvnrnd(x_est,P,100);
clf
ax = axes;
hold(ax,'on')
grid(ax,'on')
plot(ax,X(:,1),X(:,2),'+')
axis(ax,'equal')
```



ความน่าจะเป็นและสัญญาณสุ่ม (Probability & Random Signal)

ตอนที่ 8 : สัญญาณสุ่ม (Random Signal)

เราได้ศึกษาเกี่ยวกับสัญญาณซึ่งเป็นค่าที่แปรเปลี่ยนไปตามเวลา(หรือปริภูมิ) ซึ่งเราจะกำหนดให้ตัวแปร t เป็นตัวแปรอิสระที่บ่งบอกถึงเวลา **หากสัญญาณนั้นมีความไม่แน่นอน เราจะเรียกสัญญาณดังกล่าวว่า สัญญาณสุ่ม (random signal)**

หากเรามองในมุมมองของสัญญาณที่เกิดจากการทดลอง สัญญาณสุ่มคือสัญญาณที่ไม่เหมือนเดิมเมื่อทำการทดลองซ้ำ

การจำลองข้อมูลการสุ่มของสัญญาณสุ่ม

กำหนดให้เราสังเคราะห์ข้อมูลเสมือนของสัญญาณสุ่มขึ้นมา

โดยที่สัญญาณสุ่มจะมีทั้งหมด n ช่อง/ตัว

$$\mathbf{X}[k] : \mathbf{Z} \to \Re^n$$

และแต่ละช่องจะถูกเก็บค่าในช่วงเวลาทั้งหมด N ค่า (โดยสุ่มตัวอย่าง หรือ sample)

$$\bar{\mathbf{X}} = [\mathbf{X}[1] \quad \cdots \quad \mathbf{X}[N]] = \begin{bmatrix} X_1[1] & \cdots & X_1[N] \\ \vdots & \ddots & \vdots \\ X_n[1] & \cdots & X_n[N] \end{bmatrix}$$

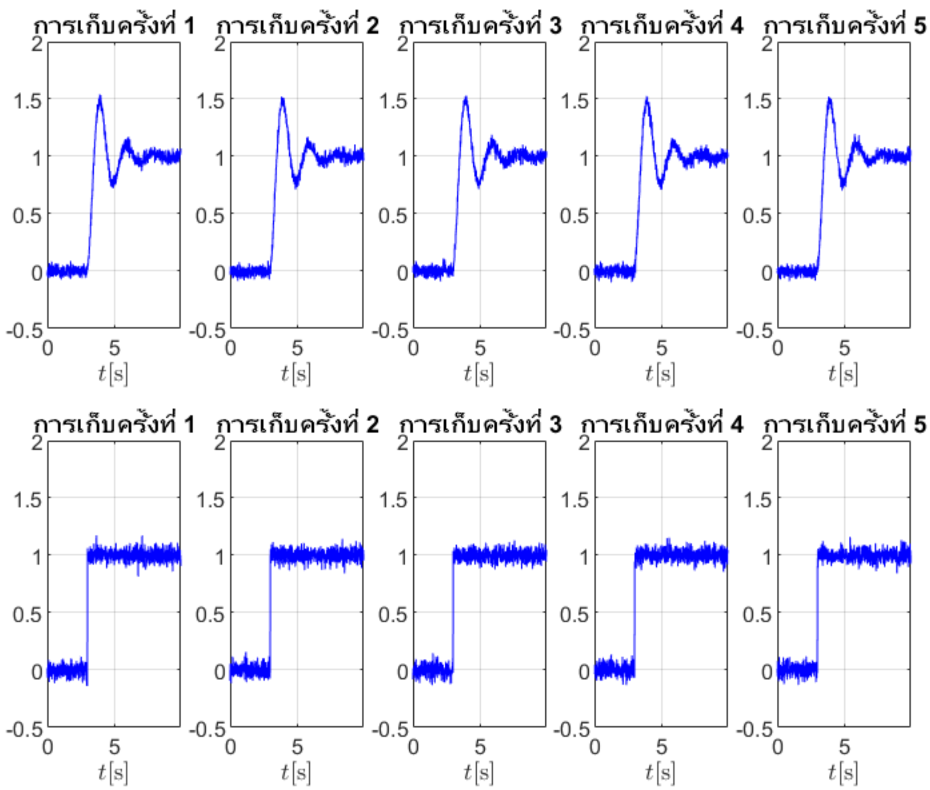
และเราเก็บค่าทั้งหมด M รอบ

เราสามารถเก็บข้อมูลดังกล่าวในรูปของ array หลายมิติได้ โดยที่เราจะกำหนดให้ มิติที่หนึ่งคือเวลา (1 ถึง N) มิติที่สองคือตัวแปร (1 ถึง n) และมิติที่สามคือครั้งในการเก็บข้อมูล (1 ถึง M) ซึ่งเราเขียนเป็น pseudo-code คร่าวๆได้ดังนี้

```
t = (0:dt:dt*(N-1));
dataset = zeros(N,n,M);
for k = 1:M
    dataset(:,:,k) = run_experiment();
end
```

```
% การจำลองข้อมูล (ไม่ใช่ข้อมูลจริง)
N = 1000;
f_s = 100;
f = 0.5;
w = 2*pi*f;
dt = 1/f_s;
t = (0:dt:dt*(N-1));
y_1 = (1-exp(-0.75*(t-3)).*cos(w*(t-3))).*(t>3);
y_2 = (t>3);
n = 2;
M = 5;
dataset = zeros(N,n,M);
var = [0.001;0.002];
covar = 2*sqrt(abs(prod(var)))*(rand-1/2);
P_sim = diag(var)+[0 covar; covar 0];
for k = 1:M
    dataset(:,:,k) = [y_1' y_2']+mvnrnd([0;0],P_sim,N);
end
```

```
clf
for j = 1:n
    for k = 1:M
        ax = subplot(n,M,M*(j-1)+k);
        plot(ax,t,dataset(:,j,k),'b')
        axis(ax,[min(t) max(t) -0.5 2])
        grid(ax,'on')
        xlabel('$t$[s]','Interpreter',"latex")
        title(ax,sprintf('การเก็บครั้งที่ %d',k))
    end
end
```



ค่าเฉลี่ย

เมื่อเราทดลองและเก็บผลหลายครั้ง เราต้องมั่นใจว่าการทดลองแต่ละครั้งนั้นถูกควบคุมให้เงื่อนไขต่างๆเป็นเหมือนกัน ผลลัพธ์ที่ได้มานั้นจะมีความแปรปรวนในแต่ละรอบการทดลอง **เราสามารถหาค่าเฉลี่ยของการทดลองได้โดยการนำผลของการทดลองแต่ละครั้งมาหารเฉลี่ยกัน แต่เราจะไม่นำค่าในช่วงเวลามาหารเฉลี่ยกัน หรือกล่าวอีกอย่างคือ เราจะไม่นำค่าที่มีเวลาเป็นคนละเวลามาหารเฉลี่ยรวมกัน**

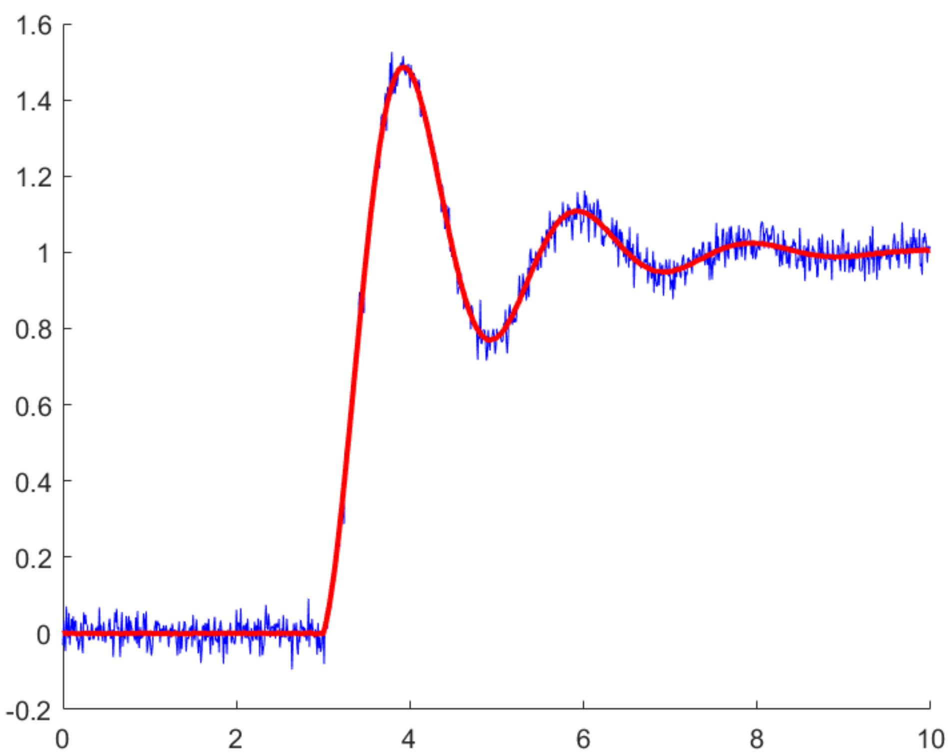
$$\mu_{\mathbf{X}}[k] = E\{\boldsymbol{X}[k]\} = \frac{1}{M}\sum_{j=1}^{M}\{\boldsymbol{X}^{(j)}[k]\}$$

โดยที่ $\boldsymbol{X}^{(j)}[k]$ คือค่าของตัวแปรสุ่มทุกตัว ณ การทดลองที่ j และเวลาที่ k

เราสามารถใช้ฟังก์ชัน mean ในการหาค่าเฉลี่ยโดยที่เราจะเฉลี่ยข้อมูลในมิติที่สาม เราสามารถคำนวณและกราฟค่าเฉลี่ยและข้อมูลได้ดังต่อไปนี้

```
N = 1000;
```

```
f_s = 100;
f = 0.5;
w = 2*pi*f;
dt = 1/f_s;
t = (0:dt:dt*(N-1));
y_1 = (1-exp(-0.75*(t-3)).*cos(w*(t-3))).*(t>3);
y_2 = (t>3);
n = 2;
M = 1000;
dataset = zeros(N,n,M);
var = [0.001;0.002];
covar = 2*sqrt(abs(prod(var)))*(rand-1/2);
P_sim = diag(var)+[0 covar; covar 0];
for k = 1:M
    dataset(:,:,k) = [y_1' y_2']+mvnrnd([0;0],P_sim,N);
end
mu = mean(dataset,3);
clf;
ax = axes;
hold(ax,'on')
plot(ax,t,dataset(:,1,1),'b')
plot(ax,t,mu(:,1),'r','LineWidth',2)
```



เมตริกซ์ค่าแปรปรวนร่วมเกี่ยว

เช่นเดียวกับการหาค่าเฉลี่ย **เราสามารถคำนวณหาเมตริกซ์ค่าแปรปรวนร่วมเกี่ยวของสัญญาณได้ โดยคำนวณหาเมตริกซ์ในแต่ละเวลา** k ผลที่ได้คือเมตริกซ์ $\mathbf{K}_\mathbf{X}[k]$ ทั้งหมด N ตัว

$$\mathbf{K_X}[k] = \mathrm{Cov}\{\boldsymbol{X}[k], \boldsymbol{X}[k]\}$$

หากในทุกเวลา k เมตริกซ์แต่ละตัวมีค่าที่ใกล้เคียงกัน เราสามารถประมาณได้ว่าคุณลักษณะทางสถิติของทั้งสัญญาณนั้นไม่แปรเปลี่ยน หรือเรียกอีกอย่างได้ว่า สัญญาณนั้นเป็นสัญญาณนิ่ง (stationary signal)

เราสามารถประมาณได้ว่าเมตริกซ์ค่าแปรปรวนร่วมเกี่ยวได้ดังต่อไปนี้

$$\mathbf{K_X} \approx \frac{1}{N}\sum_{k=1}^{N} \mathbf{K_X}[k]$$

เราสามารถเขียนโปรแกรมจำลองการทดลองได้ดังต่อไปนี้

```
N = 1000;
f_s = 100;
f = 0.5;
w = 2*pi*f;
dt = 1/f_s;
t = (0:dt:dt*(N-1));
y_1 = (1-exp(-0.75*(t-3)).*cos(w*(t-3))).*(t>3);
y_2 = (t>3);
n = 2;
M = 1000;
dataset = zeros(N,n,M);
var = [0.001;0.002];
covar = 2*sqrt(abs(prod(var)))*(rand-1/2);
P_sim = diag(var)+[0 covar; covar 0];
for k = 1:M
    dataset(:,:,k) = [y_1' y_2']+mvnrnd([0;0],P_sim,N);
end
mu = mean(dataset,3);
K_XX = zeros(n);
for j = 1:N
    % for each t, compute covriance matrix
    K_XX = K_XX + cov(permute(dataset(j,:,:),[3,2,1]));
end
K_XX = K_XX/N
```

```
K_XX = 2×2
    0.0010    0.0002
    0.0002    0.0020
```

เราสามารถนำเมตริกซ์ดังกล่าวเพื่ออธิบายความไม่แน่นอนของค่าที่วัดได้จากเซนเซอร์และนำไปใช้ต่อในการทำตัวกรองคาลมาน (Kalman Filter)