

$$P(\gamma_2, \gamma_1, \gamma_0 | \alpha) = P(\gamma_2 | \gamma_1) \cdot P(\gamma_1, \gamma_0) \cdot P(\gamma_0 | \alpha)$$

ឱ្យ P ជា normal distribution រួច $= \frac{1}{\sqrt{2\pi b^2}} \cdot e^{-\frac{(x-\mu)^2}{2b^2}}$

ដោយឡែក យើងដឹងថា $\gamma_2 \sim N(\alpha, \gamma_1, b)$ គឺជា distribution របស់ γ_2 ផងដែរ

$$\text{គឺជា } P(\gamma_2 | \gamma_1) = \frac{1}{\sqrt{2\pi b^2}} \cdot e^{-\frac{(\gamma_2 - \alpha\gamma_1)^2}{2b^2}}$$

$$P(\gamma_1 | \gamma_0) = \frac{1}{\sqrt{2\pi b^2}} \cdot e^{-\frac{(\gamma_1 - \alpha\gamma_0)^2}{2b^2}}$$

$P(\gamma_0 | \alpha)$ ឱ្យ α ជា parameter យើងដឹងថា γ_0 គឺជា distribution របស់ α រួច

$$\text{គឺជា } P(\gamma_0 | \alpha) = P(\gamma_0) = \frac{1}{\sqrt{2\pi k^2}} \cdot e^{-\frac{(\gamma_0 - 0)^2}{2k^2}}$$

$$\begin{aligned} \text{គឺជា } \ln(P(\gamma_2, \gamma_1, \gamma_0 | \alpha)) &= \ln\left(\frac{1}{\sqrt{2\pi b^2}}\right) - \frac{(\gamma_2 - \alpha\gamma_1)^2}{2b^2} + \ln\left(\frac{1}{\sqrt{2\pi b^2}}\right) - \frac{(\gamma_1 - \alpha\gamma_0)^2}{2b^2} \\ &\quad + \ln\left(\frac{1}{\sqrt{2\pi k^2}}\right) - \frac{(\gamma_0)^2}{2k^2} \end{aligned}$$

$$0 = 0 - \frac{2(\gamma_2 - \alpha\gamma_1) \cdot (-\gamma_1)}{2b^2} - \frac{2(\gamma_1 - \alpha\gamma_0) \cdot (-\gamma_0)}{2b^2}$$

$$0 = \frac{\gamma_2\gamma_1 - \alpha\gamma_1^2}{b} + \frac{\gamma_1\gamma_0 - \alpha\gamma_0^2}{b}$$

$$\alpha(\gamma_1^2 + \gamma_0^2) = \gamma_2\gamma_1 + \gamma_1\gamma_0$$

$$\alpha = \frac{(\gamma_2\gamma_1 + \gamma_1\gamma_0)}{\gamma_1^2 + \gamma_0^2}$$

$$\hat{\alpha}_{MLE} = \frac{\gamma_1(\gamma_2 + \gamma_0)}{\gamma_1^2 + \gamma_0^2} \quad \text{X}$$

4.

```
print('''
ข้อ 4-1 ตอบ ไม่ได้เพราะค่า average number ของ visitors เมื่อ block + new channel ปฏิเสธ H0 จริง การเพิ่มขึ้นของ Visitors อาจเกิดจากการ Block ของ government ก็ได้ว่าจะไม่เกิดจาก
การที่เพิ่ม new channel อย่างเดียว
''')
print('''
ข้อ 4-2 ตอบ จากโจทย์กำหนดว่าการที่ user จะเข้าช่องทางไหนเป็นอิสระต่อกันนั่นหมายความว่า จะ Block หรือไม่ block นั้นไม่เกี่ยวข้องกับการเข้าช่องทางใหม่หรือเก่าทำให้ ดังนั้น H1:  $x_1 > x_2$  กับ H1:  $x_3 > x_0$ 
คือสมมติฐานเดียวกัน แต่ Assume ว่า Hamtaro กำหนด Significant level เป็น 0.05 (โจทย์ไม่ได้บอก) ก็จะได้ว่า P value > 0.05 ซึ่งไม่ปฏิเสธ H0 เลยต้องตอบว่าไม่สามารถสรุปได้ว่าการเพิ่ม Channel
แล้วจะเพิ่ม average visitors อย่างมีนัยสำคัญ
''')
print('''
ข้อ 4-3 ตอบ จากข้อ 1 ถ้าเรา Assume significant level (alpha) เป็น 0.05 ก็จะได้ว่า H1 เป็นจริงหรือก็คือปฏิเสธ H0 ซึ่งทำให้เราได้ข้อมูลเพิ่มว่า การเพิ่มขึ้นของ Channel และ Blocking ทำให้เพิ่ม
average visitors อย่างมีนัยสำคัญ และจากข้อสองเราได้ว่า การเพิ่ม new channel ไม่สามารถสรุปได้ว่า average visitors เพิ่มขึ้น ทำให้เราการ setup 1. and 2. สามารถสรุปปัญหาได้ว่า การเพิ่มขึ้นของ
Channel ไม่ได้เพิ่ม average visitors อย่างมีนัยสำคัญ
''')
```

ข้อ 4-1 ตอบ ไม่ได้เพราะค่า average number ของ visitors เมื่อ block + new channel ปฏิเสธ H0 จริง การเพิ่มขึ้นของ Visitors อาจเกิดจากการ Block ของ government ก็ได้ว่าจะไม่เกิดจาก
การที่เพิ่ม new channel อย่างเดียว

ข้อ 4-2 ตอบ จากโจทย์กำหนดว่าการที่ user จะเข้าช่องทางไหนเป็นอิสระต่อกันนั่นหมายความว่า จะ Block หรือไม่ block นั้นไม่เกี่ยวข้องกับการเข้าช่องทางใหม่หรือเก่าทำให้ ดังนั้น H1: $x_1 > x_2$ กับ H1: $x_3 > x_0$
คือสมมติฐานเดียวกัน แต่ Assume ว่า Hamtaro กำหนด Significant level เป็น 0.05 (โจทย์ไม่ได้บอก) ก็จะได้ว่า P value > 0.05 ซึ่งไม่ปฏิเสธ H0 เลยต้องตอบว่าไม่สามารถสรุปได้ว่าการเพิ่ม Channel
แล้วจะเพิ่ม average visitors อย่างมีนัยสำคัญ

ข้อ 4-3 ตอบ จากข้อ 1 ถ้าเรา Assume significant level (alpha) เป็น 0.05 ก็จะได้ว่า H1 เป็นจริงหรือก็คือปฏิเสธ H0 ซึ่งทำให้เราได้ข้อมูลเพิ่มว่า การเพิ่มขึ้นของ Channel และ Blocking ทำให้เพิ่ม
average visitors อย่างมีนัยสำคัญ และจากข้อสองเราได้ว่า การเพิ่ม new channel ไม่สามารถสรุปได้ว่า average visitors เพิ่มขึ้น ทำให้เราการ setup 1. and 2. สามารถสรุปปัญหาได้ว่า การเพิ่มขึ้นของ
Channel ไม่ได้เพิ่ม average visitors อย่างมีนัยสำคัญ

5.1 และ 5.2

```
print('''
ข้อ 1 ตอบ ในฐานะผู้เล่น H0 จะเป็น Probability >= 1/6 H1 จะได้ว่า probability < 1/6
''')
print('''
ข้อ 2 ตอบ จาก H1 จะได้ว่า H1 เป็น one sided โดยที่ two sided จะแตกต่างจาก One sided ตรงที่ two sided เราสามารถสรุปได้ว่ามันไม่ยุติธรรมโดยไม่ได้สนใจว่า หน้าที่เราเลือกจะออกมากขึ้นหรือน้อยลง
ซึ่งมีข้อคิดตรงที่มันสรุปได้ตรงที่ว่า การทอยลูกเต๋านั้นยุติธรรมหรือไม่ยุติธรรม แต่ถ้าเป็น One sided จะทำให้เราสรุปได้แค่ เราได้ประโยชน์ (Probability > 1/6) หรือเราเสียประโยชน์ (Probability < 1/6) เท่านั้น
ซึ่งมีข้อคิดว่า ถ้าเราเลือก H1 เป็น P < 1/6 แปลว่าเรากำลังสงสัยว่า Hamtaro กำลังโกงเราอยู่
''')
```

ข้อ 1 ตอบ ในฐานะผู้เล่น H0 จะเป็น Probability >= 1/6 H1 จะได้ว่า probability < 1/6

ข้อ 2 ตอบ จาก H1 จะได้ว่า H1 เป็น one sided โดยที่ two sided จะแตกต่างจาก One sided ตรงที่ two sided เราสามารถสรุปได้ว่ามันไม่ยุติธรรมโดยไม่ได้สนใจว่า หน้าที่เราเลือกจะออกมากขึ้นหรือน้อยลง
ซึ่งมีข้อคิดตรงที่มันสรุปได้ตรงที่ว่า การทอยลูกเต๋านั้นยุติธรรมหรือไม่ยุติธรรม แต่ถ้าเป็น One sided จะทำให้เราสรุปได้แค่ เราได้ประโยชน์ (Probability > 1/6) หรือเราเสียประโยชน์ (Probability < 1/6) เท่านั้น
ซึ่งมีข้อคิดว่า ถ้าเราเลือก H1 เป็น P < 1/6 แปลว่าเรากำลังสงสัยว่า Hamtaro กำลังโกงเราอยู่

5.3, 5.4 และ 5.5

```
# 3.
prob_cdf = binom.cdf(3, 30, 1/6)
# print(prob_cdf)
if (prob_cdf <= 0.1):
    print(f'ปฏิเสธ H0 เพราะ Probability ที่ได้มีค่า:{round(prob_cdf,4)} ซึ่งน้อยกว่า 0.1')
else:
    print(f'ไม่ปฏิเสธ H0 เพราะ Probability ที่ได้มีค่า:{round(prob_cdf,4)} ซึ่งมีค่ามากกว่าเท่ากับ 0.1')
```

ไม่ปฏิเสธ H0 เพราะ Probability ที่ได้มีค่า:0.2396 ซึ่งมีค่ามากกว่าเท่ากับ 0.1

+ ได้ค้ + ข้อความ

```
# 4. หา k ที่มากที่สุดที่ทำให้ prob <= 0.1
k = 1
while (binom.cdf(k,200,1/6) <= 0.1):
    k+=1
print(f'Rejection region มีค่าเป็น {k-1}')
```

Rejection region มีค่าเป็น 26

```
# 5.
k = 1
val = norm.cdf(k, (1/6)*200,np.sqrt(200*(1/6)*(5/6)))
while (val <= 0.1):
    k+=1
val = norm.cdf(k, (1/6)*200,np.sqrt(200*(1/6)*(5/6)))
print(f'Rejection region มีค่าเป็น {k-1}')
```

Rejection region มีค่าเป็น 26

5.6

```
# 6.  
# find k  
k = 1  
while (binom.cdf(k,200,1/6) <= 0.01):  
    k+=1  
k-=1  
# find p min  
lo, hi = 1e-12, 1 - 1e-12  
while(True):  
    mid = (lo + hi) / 2  
    p_mid = binom.cdf(k, 200, mid)  
    if abs(p_mid - 0.05) < 1e-10:  
        break  
    if p_mid > 0.05:  
        lo = mid  
    else:  
        hi = mid  
print(round((lo+hi)/2,3))
```

```
0.148
```

5.7

```
# 7.  
# find k  
k = 1  
while (binom.cdf(k,200,1/6) <= 0.01):  
    k+=1  
k-=1  
# find p min  
lo, hi = 1e-12, 1 - 1e-12  
while(True):  
    mid = (lo + hi) / 2  
    p_mid = binom.cdf(k, 200, mid)  
    if abs(p_mid - 0.01) < 1e-10:  
        break  
    if p_mid > 0.01:  
        lo = mid  
    else:  
        hi = mid  
print(round((lo+hi)/2,5))
```

```
⇒ 0.16602
```

7.1 และ 7.2

```
# 1.
print('
ให้ x2: เป็น number of goods product ของ new machine และให้ u2 เป็นค่าเฉลี่ยของ good product ใน factories 4 factories ที่ใช้ new machine
จะได้ H0: u2 <= 5000
และ H1: u2 > 5000 (เพื่อทดสอบว่าจะปฏิเสธ H0 นี้จะได้ความได้ว่าดีขึ้นหรือป่าว)
')
```

ให้ x2: เป็น number of goods product ของ new machine และให้ u2 เป็นค่าเฉลี่ยของ good product ใน factories 4 factories ที่ใช้ new machine
จะได้ H0: u2 <= 5000
และ H1: u2 > 5000 (เพื่อทดสอบว่าจะปฏิเสธ H0 นี้จะได้ความได้ว่าดีขึ้นหรือป่าว)

```
# 2.
x_bar = (sum(fac_0)+sum(fac_1)+sum(fac_2)+sum(fac_3))/120
z_val = (x_bar - 5000)/(20/np.sqrt(120))
prob = norm.cdf(z_val)
# จากข้อนี้เป็นทดสอบ Right tail เพราะ H1 เป็นเครื่องหมายมากกว่าทำให้จะปฏิเสธ H0 เมื่อ prob มากกว่า 1-significant level
if (prob > (1-0.05)):
    print(f"สามารถสรุปได้ว่าปฏิเสธ H0 ทำให้สรุปได้ว่าการใช้เครื่องจักรใหม่สามารถเพิ่ม productivity โดยรวมได้อย่างมีนัยสำคัญเพราะ {prob} > {1-0.05}")
else:
    print(f"ไม่สามารถสรุปได้ว่าปฏิเสธ H0 เพราะ {prob} < {1-0.05} ทำให้สรุปไม่ได้ว่าเครื่องจักรใหม่สามารถเพิ่ม productivity ได้อย่างมีนัยสำคัญ")

สามารถสรุปได้ว่าปฏิเสธ H0 ทำให้สรุปได้ว่าการใช้เครื่องจักรใหม่สามารถเพิ่ม productivity โดยรวมได้อย่างมีนัยสำคัญเพราะ 0.9999886459853142 > 0.95
```

7.3

```
# 3.
x = [x1,x2,x3,x4] = sum(fac_0)/30,sum(fac_1)/30,sum(fac_2)/30,sum(fac_3)/30
z = [z1,z2,z3,z4] = (x1 - 5000)/(20/np.sqrt(30)),(x2 - 5000)/(20/np.sqrt(30)),(x3 - 5000)/(20/np.sqrt(30)),(x4 - 5000)/(20/np.sqrt(30))
for i in range(len(z)):
    prob = norm.cdf(z[i])
    if (prob > (1-0.05)):
        print(f"สามารถสรุปได้ว่าปฏิเสธ H0 ทำให้สรุปได้ว่าการใช้เครื่องจักรใหม่สามารถเพิ่ม productivity ของ {fac[i]} ได้อย่างมีนัยสำคัญเพราะ {prob} > {1-0.05}")
    else:
        print(f"ไม่สามารถสรุปได้ว่าปฏิเสธ H0 เพราะ {prob} < {1-0.05} ทำให้สรุปไม่ได้ว่าเครื่องจักรใหม่สามารถเพิ่ม productivity ของ {fac[i]} ได้อย่างมีนัยสำคัญ")

สามารถสรุปได้ว่าปฏิเสธ H0 ทำให้สรุปได้ว่าการใช้เครื่องจักรใหม่สามารถเพิ่ม productivity ของ fac_0 ได้อย่างมีนัยสำคัญเพราะ 0.9847931471862857 > 0.95
สามารถสรุปได้ว่าปฏิเสธ H0 ทำให้สรุปได้ว่าการใช้เครื่องจักรใหม่สามารถเพิ่ม productivity ของ fac_1 ได้อย่างมีนัยสำคัญเพราะ 0.9988717027389801 > 0.95
สามารถสรุปได้ว่าปฏิเสธ H0 ทำให้สรุปได้ว่าการใช้เครื่องจักรใหม่สามารถเพิ่ม productivity ของ fac_2 ได้อย่างมีนัยสำคัญเพราะ 0.9596631595129791 > 0.95
ไม่สามารถสรุปได้ว่าปฏิเสธ H0 เพราะ 0.9341265256796157 < 0.95 ทำให้สรุปไม่ได้ว่าเครื่องจักรใหม่สามารถเพิ่ม productivity ของ fac_3 ได้อย่างมีนัยสำคัญ
```

7.4.1

```
# 4.1
print('
ให้ x2: เป็น number of goods product ของ new machine และให้ u2 เป็นค่าเฉลี่ยของ good product ใน factories 4 factories ที่ใช้ new machine
จะได้ H0: u2 <= 5000
และ H1: u2 > 5000 (เพื่อทดสอบว่าจะปฏิเสธ H0 นี้จะได้ความได้ว่าดีขึ้นหรือป่าว)
')
```

ให้ x2: เป็น number of goods product ของ new machine และให้ u2 เป็นค่าเฉลี่ยของ good product ใน factories 4 factories ที่ใช้ new machine
จะได้ H0: u2 <= 5000
และ H1: u2 > 5000 (เพื่อทดสอบว่าจะปฏิเสธ H0 นี้จะได้ความได้ว่าดีขึ้นหรือป่าว)

7.4.2

```
# 4.2
con_arr = np.concatenate((fac_0,fac_1,fac_2,fac_3))
x_bar = con_arr.mean()
variance = (1/(n-1))*(sum((con_arr[i] - x_bar)**2 for i in range(len(con_arr))))
s = np.sqrt(variance)
t_x = (x_bar - u)/(s/np.sqrt(n))
# degree of freedom be n-1 = 119
prob_cdf = t.cdf(t_x,119)
if (prob_cdf > (1-0.05)):
    print("สามารถสรุปได้ว่าปฏิเสธ H0 ทำให้สรุปได้ว่าการใช้เครื่องจักรใหม่สามารถเพิ่ม productivity โดยรวมได้อย่างมีนัยสำคัญ")
else:
    print(f"ไม่สามารถสรุปได้ว่าปฏิเสธ H0 เพราะ {prob_cdf} < {1-0.05} ทำให้สรุปไม่ได้ว่าเครื่องจักรใหม่สามารถเพิ่ม productivity ได้อย่างมีนัยสำคัญ")
```

สามารถสรุปได้ว่าปฏิเสธ H0 ทำให้สรุปได้ว่าการใช้เครื่องจักรใหม่สามารถเพิ่ม productivity โดยรวมได้อย่างมีนัยสำคัญ

7.4.3

```
# 4.3
for i in range (len(fac)):
    x_bar = l[i].mean()
    variance = (1/(n-1))*(sum((l[i][j] - x_bar)**2 for j in range(len(l[i]))))
    s = np.sqrt(variance)
    t_x = (x_bar - u)/(s/np.sqrt(n))
    # degree of freedom be n-1 = 1
    prob_cdf = t.cdf(t_x,119)
    if (prob_cdf > (1-0.05)):
        print(f"สามารถสรุปได้ว่าปฏิเสธ H0 ทำให้สรุปได้ว่าการใช้เครื่องจักรใหม่สามารถเพิ่ม productivity ของ {fac[i]} ได้อย่างมีนัยสำคัญเพราะ prob {prob_cdf} > {1-0.05}")
    else:
        print(f"ไม่สามารถสรุปได้ว่าปฏิเสธ H0 เพราะ {prob_cdf} < {1-0.05} ทำให้สรุปไม่ได้ว่าเครื่องจักรใหม่สามารถเพิ่ม productivity ของ {fac[i]} ได้อย่างมีนัยสำคัญ")
```

สามารถสรุปได้ว่าปฏิเสธ H0 ทำให้สรุปได้ว่าการใช้เครื่องจักรใหม่สามารถเพิ่ม productivity ของ fac_0 ได้อย่างมีนัยสำคัญเพราะ prob 0.9999999999999999 > 0.95
สามารถสรุปได้ว่าปฏิเสธ H0 ทำให้สรุปได้ว่าการใช้เครื่องจักรใหม่สามารถเพิ่ม productivity ของ fac_1 ได้อย่างมีนัยสำคัญเพราะ prob 1.0 > 0.95
สามารถสรุปได้ว่าปฏิเสธ H0 ทำให้สรุปได้ว่าการใช้เครื่องจักรใหม่สามารถเพิ่ม productivity ของ fac_2 ได้อย่างมีนัยสำคัญเพราะ prob 0.9999999999833026 > 0.95
สามารถสรุปได้ว่าปฏิเสธ H0 ทำให้สรุปได้ว่าการใช้เครื่องจักรใหม่สามารถเพิ่ม productivity ของ fac_3 ได้อย่างมีนัยสำคัญเพราะ prob 0.9999999975864656 > 0.95