01

以蒙地卡羅實驗驗證 J-B檢定統計量

黃暐宸

甚麼是統計量G1, G2, G3?

統計量G1
(偏度的調整值)

統計量G1

偏度的定義與意義

→ 是一種衡量數據分佈偏斜方向和程度的統計量

G1是樣本偏度的調整值

→ $$G_1 = \sqrt{\frac{n}{6}}\,\hat{s}$$

→ ŝ 是偏態係數的估計值

統計量G2
(峰度的調整值)

統計量G2

峰度的定義與意義

→ 是描述一組數據分佈形狀的統計量，特別是該分佈平均值(頭)附近的尖度或扁平度

G2是樣本峰度的調整值

→ $$G_2 = \sqrt{\frac{n}{24}}(\hat{k} - 3)$$

→ $\hat{k}$ 是峰態係數的估計值

J-B (Jarque-Bera) 常態檢定統計量 (統計量G3)

甚麼是J-B檢定?

J–B檢定
(Jarque–Bera檢定)

定義與意義

➡ J–B檢定是對樣本數據是否具有符合常態分布的檢定

➡ 統計量G3(J–B檢定量)近似服從自由度為2的卡方分布

統計量G3計算方法

➡ $$G_3 = G_1^2 + G_2^2 = \frac{n}{6}\left(\hat{s}^2 + \frac{(\hat{k}-3)^2}{4}\right)$$

甚麼是蒙地卡羅實驗(模擬)?

- 使用電腦、程式來模擬實際狀況
- 以機率為基礎的一種計算方式
- 重複實驗

作品需要用到的模組

```
from scipy.stats import norm, skew, kurtosis, ttest_ind, chi2, t, uniform
from scipy.stats import anderson, norm, jarque_bera, normaltest, kstest, shapiro
import matplotlib.pyplot as plt
import numpy as np
```

作品目標

第一部分: 觀察 G1, G2, G3 的抽樣分配與樣本的關係。

第二部分: 以蒙地卡羅實驗驗證 J-B 檢定統計量，把樣本從小做到大，觀察他的表現，來看J-B 檢定統計量是不是好的檢定統計量，接著再與其他檢定方法比較好壞。

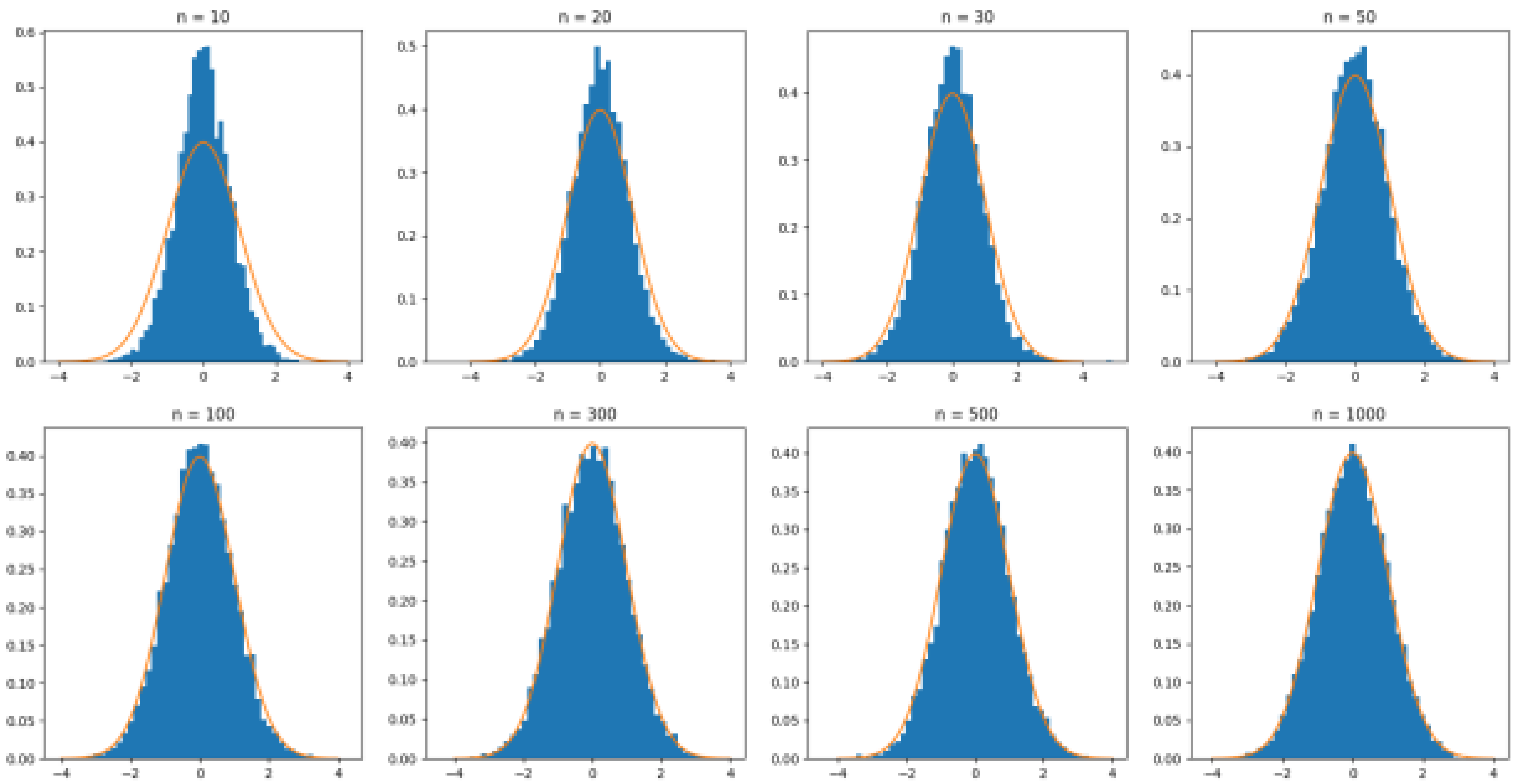
第一部分:
觀察 G1, G2, G3 的抽樣分配與樣本的關係

透過蒙地卡羅模擬驗證統計量G1服從標準常態N(0, 1)

蒙地卡羅模擬的環境設定（scenarios）：

1. 樣本數 n = 10, 20, 30, 50 , 100, 300, 500, 1000。
2. 針對每個樣本數 ，模擬次數皆為 N = 50000。
3. 繪製 n=10 與 n = 500時，統計量 G1 的直方圖與 ECDF 圖。並分別畫上對應的標準常態 PDF 與 CDF 圖。

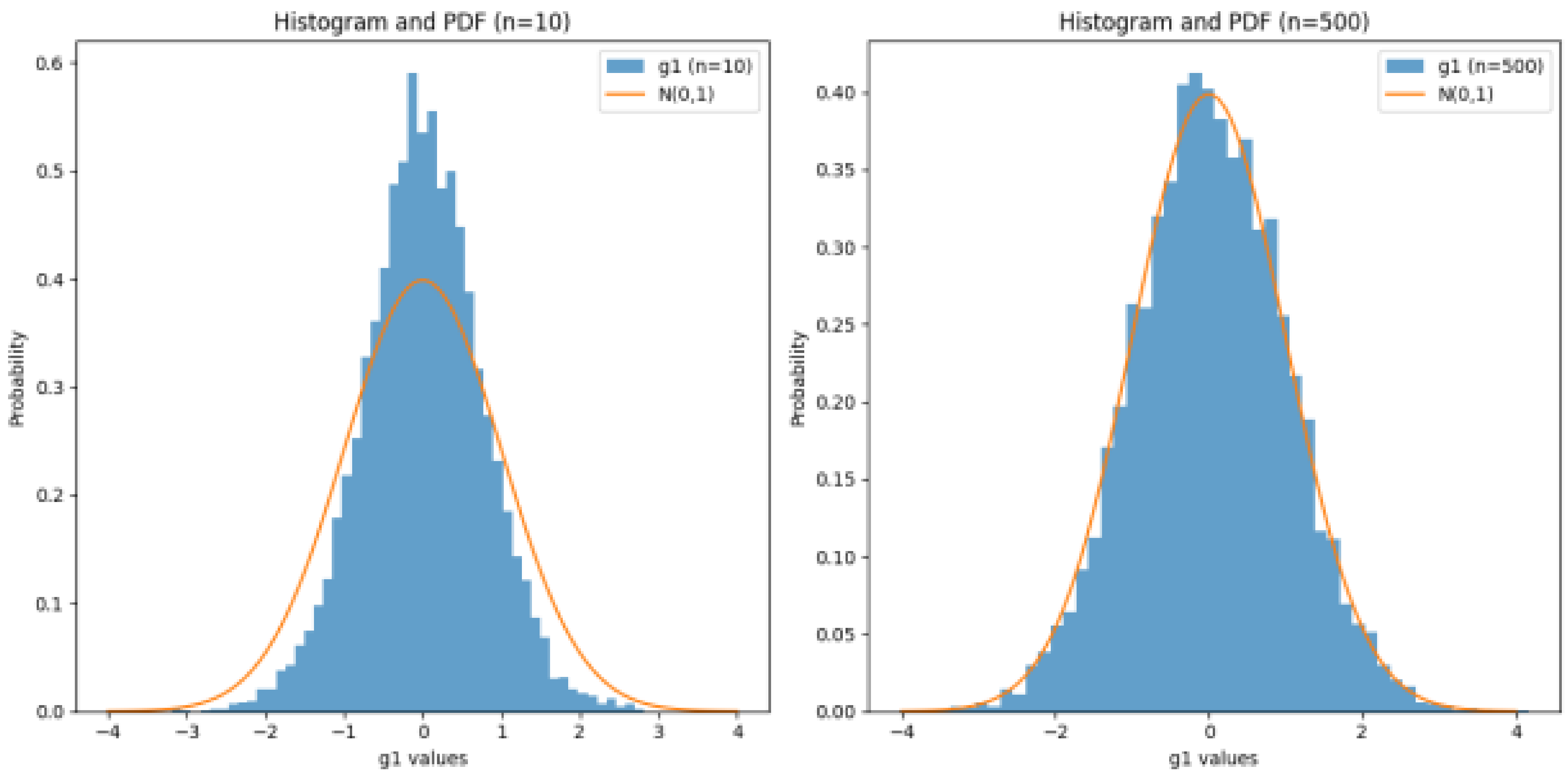
觀察檢定統計量G1的直方圖與對應的標準常態PDF



觀察檢定統計量 G1 的直方圖和對應的標準常態 PDF

樣本數 n=10 時統計量G1的直方圖不太接近標準常態PDF

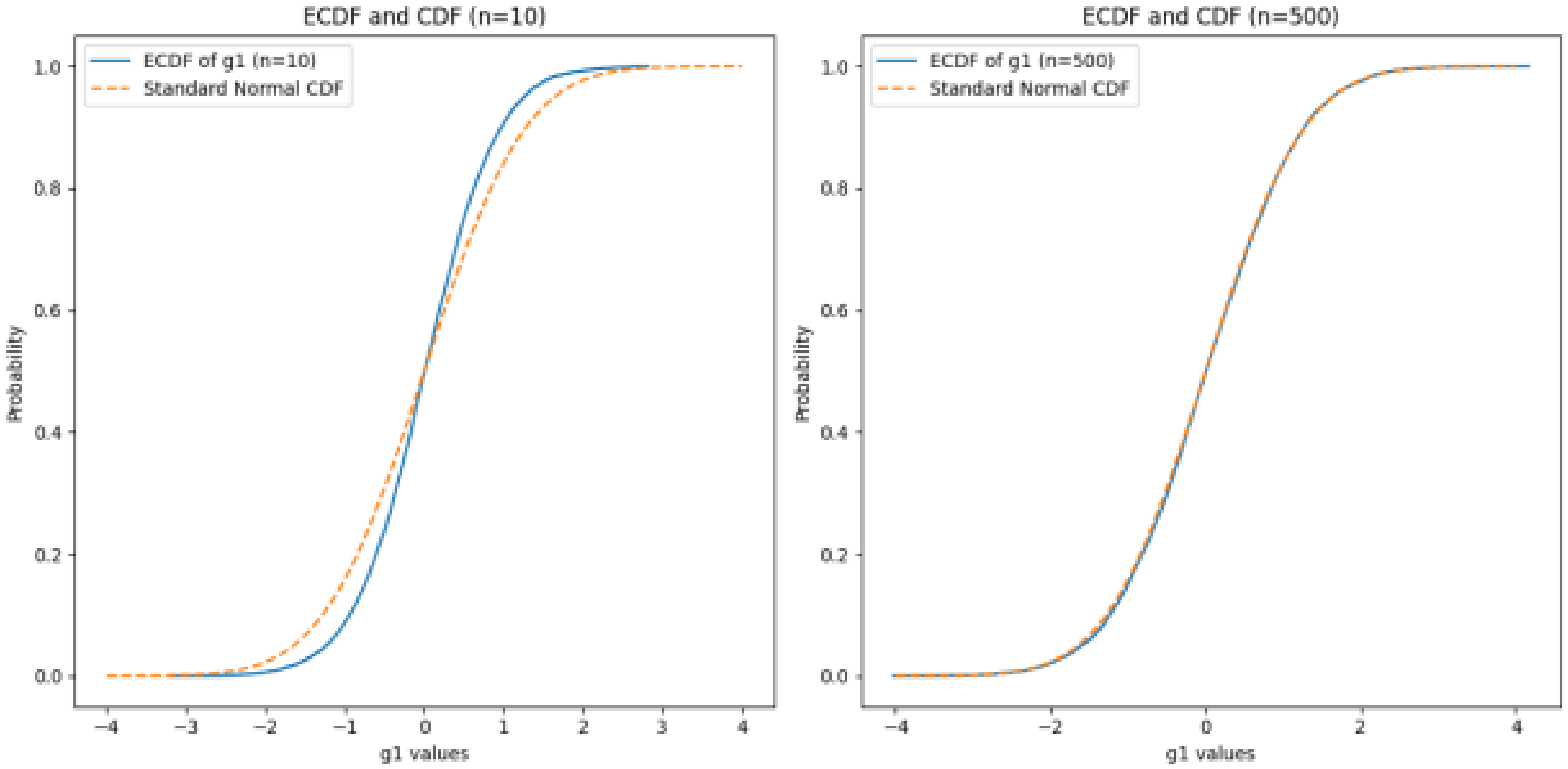
當樣本數增大到 500 時，比在樣本數10時更接近



觀察檢定統計量 G1 的 ECDF 圖和對應的標準常態 CDF

統計量的 ECDF 圖在樣本數 n = 10時，與標準常態的 CDF 圖在上升的階段差異明顯

樣本數增大到 500 時，非常接近於標準常態的 CDF 圖

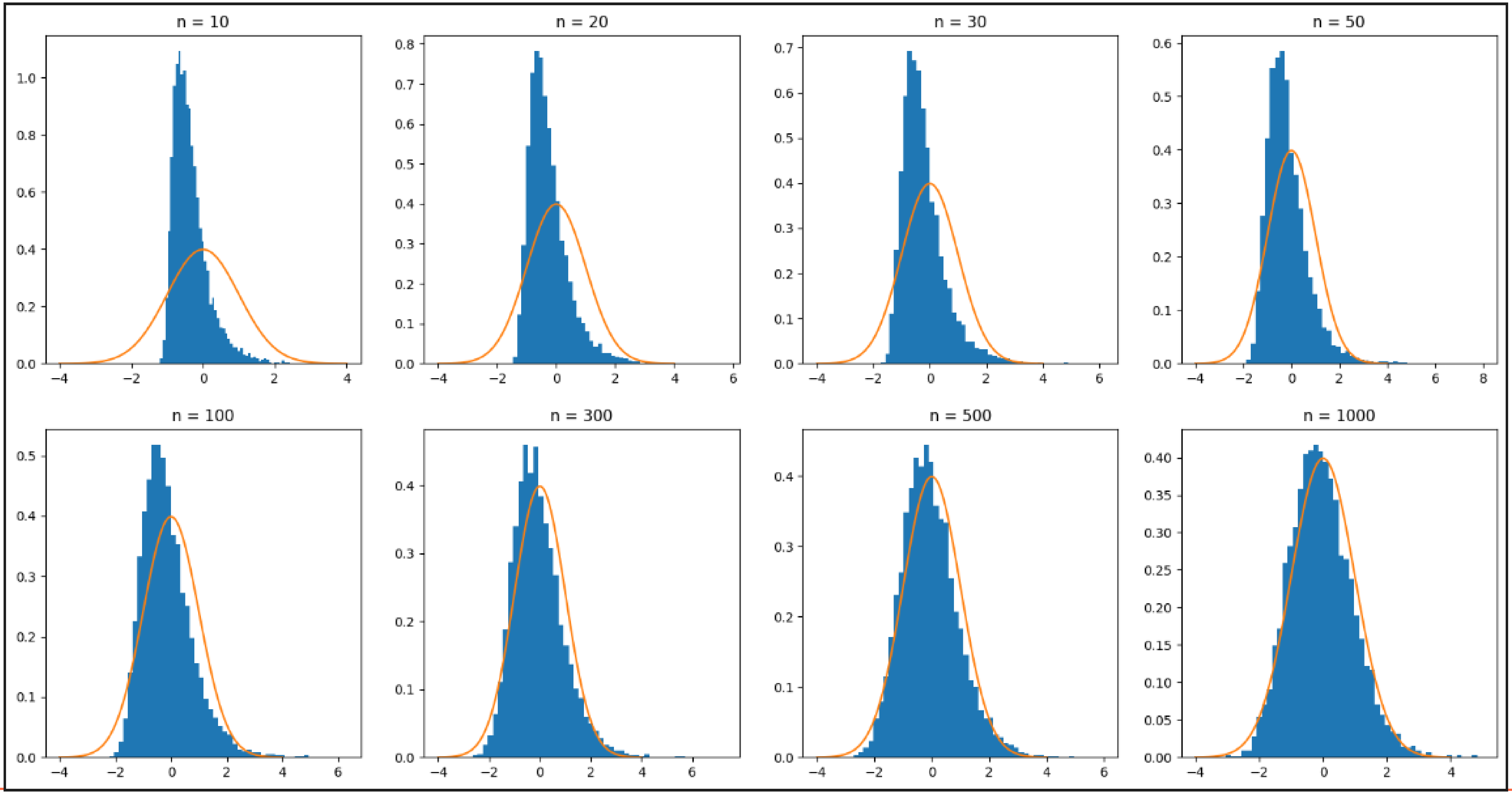


透過蒙地卡羅模擬驗證統計量G2服從標準常態N(0, 1)

蒙地卡羅模擬的環境設定（scenarios）：

1. 樣本數 n = 10, 20, 30, 50 , 100, 300, 500, 1000。
2. 針對每個樣本數 ，模擬次數皆為 N = 50000。
3. 繪製 n=10 與 n = 500時，統計量 G2的直方圖與 ECDF 圖。並分別畫上對應的標準常態 PDF 與 CDF 圖。

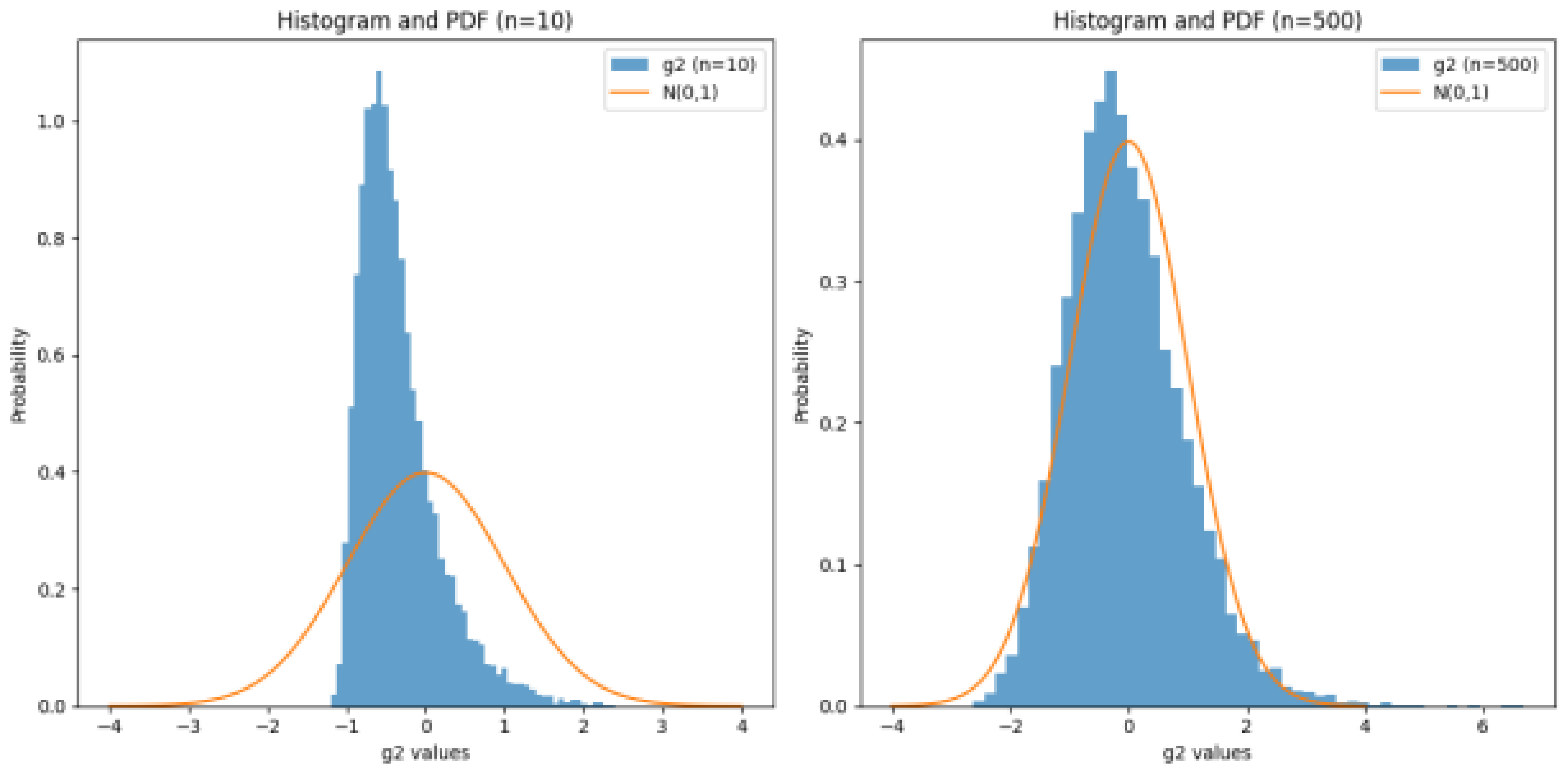
觀察檢定統計量G2的直方圖與對應的標準常態PDF



觀察檢定統計量 G2 的直方圖和對應的標準常態 PDF

樣本數 n=10 時統計量G2的直方圖不太接近標準常態PDF

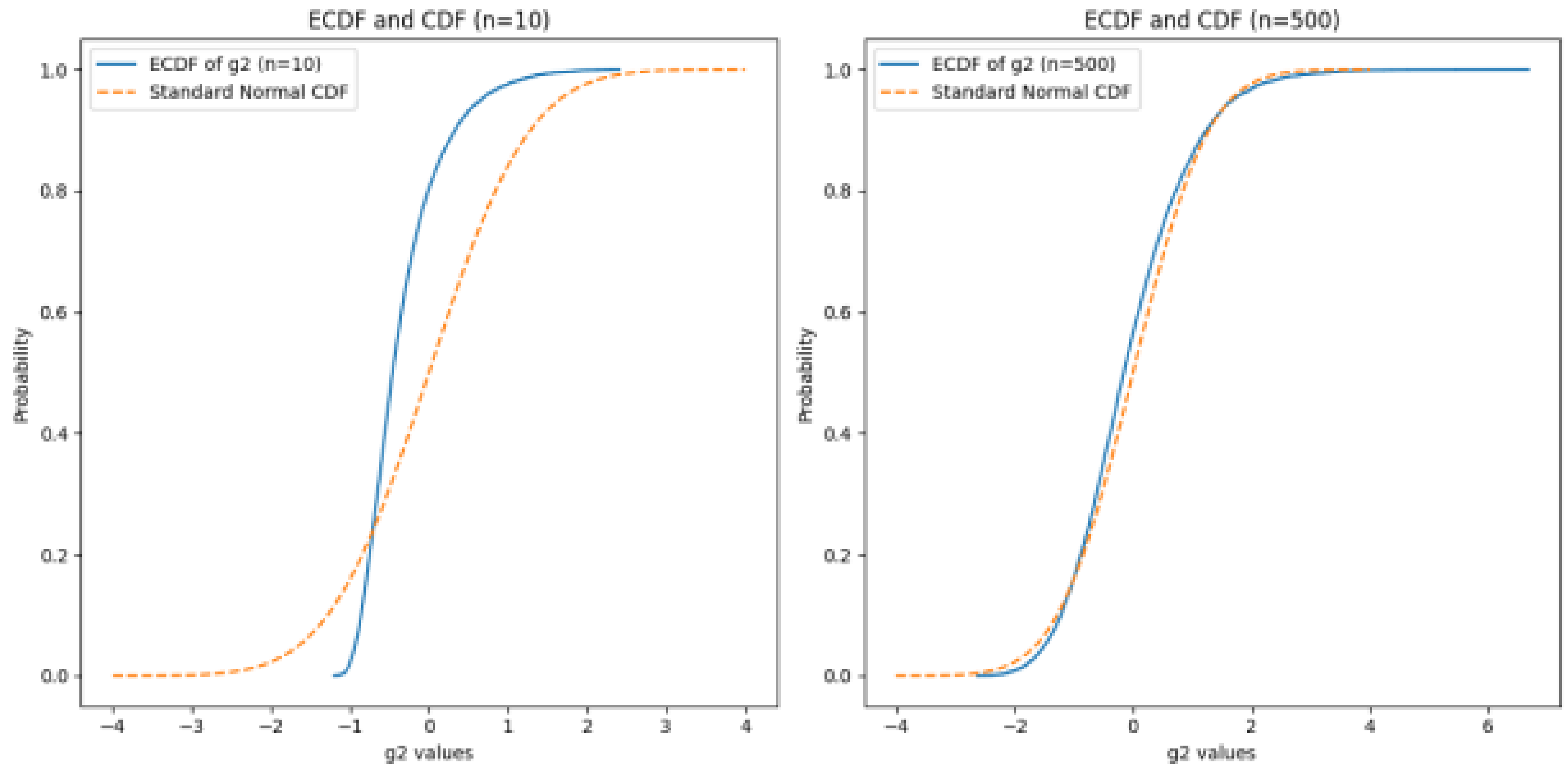
當樣本數增大到 500 時，比在樣本數10時更接近



觀察檢定統計量 G2 的 ECDF 圖和對應的標準常態 CDF

統計量的 ECDF 圖在樣本數 n = 10 時，與標準常態的 CDF 圖在上升的階段差異明顯

樣本數增大到 500 時，非常接近於標準常態的 CDF 圖

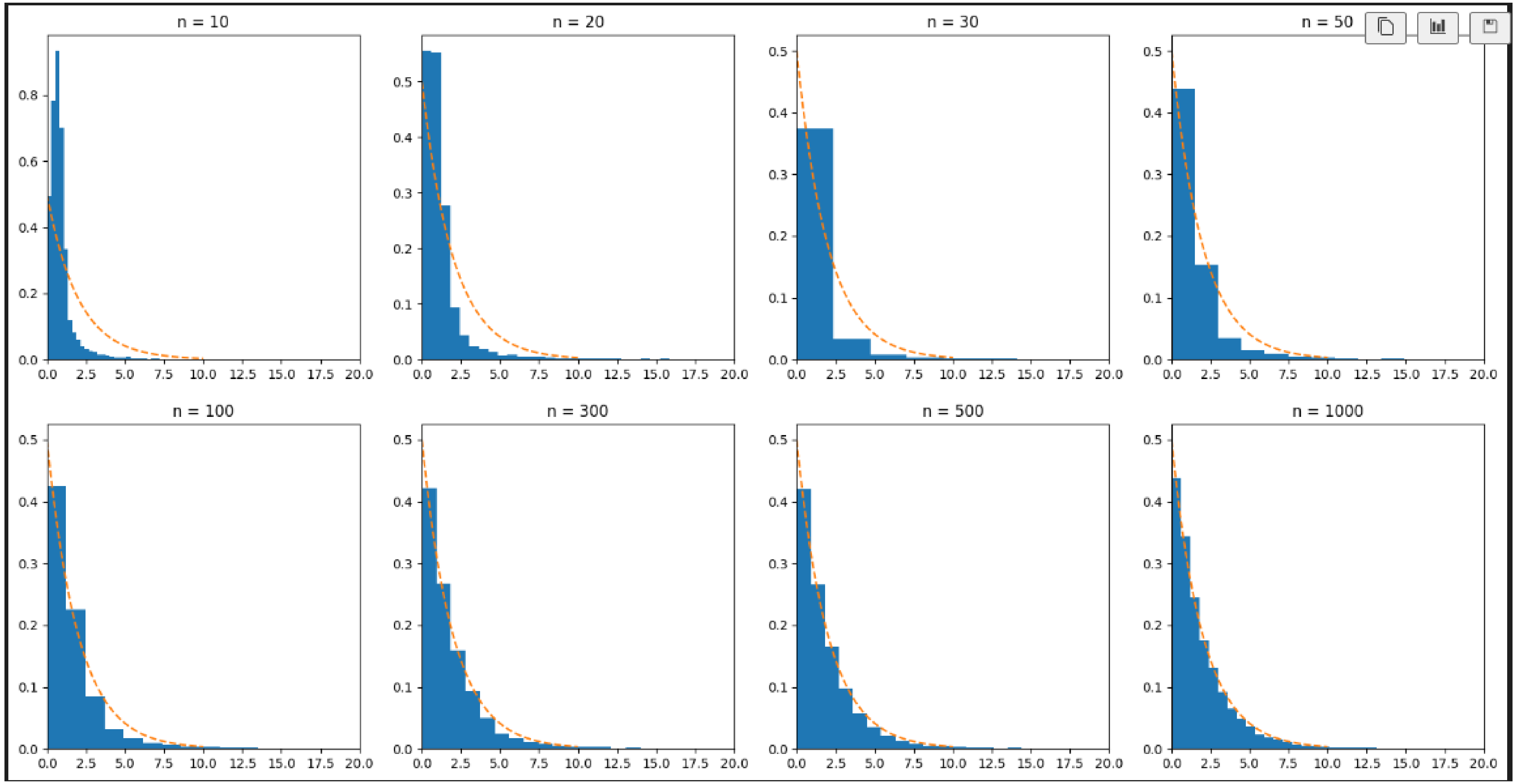


透過蒙地卡羅模擬驗證統計量G3服從卡方分配(df=2)

蒙地卡羅模擬的環境設定（scenarios）：

1. 樣本數 n = 10, 20, 30, 50 , 100, 300, 500, 1000。
2. 針對每個樣本數 ，模擬次數皆為 N = 50000。
3. 繪製 n=10 與 n = 500時，統計量 G3 的直方圖與 ECDF 圖。並分別畫上對應的卡方的 PDF 與 CDF 圖。

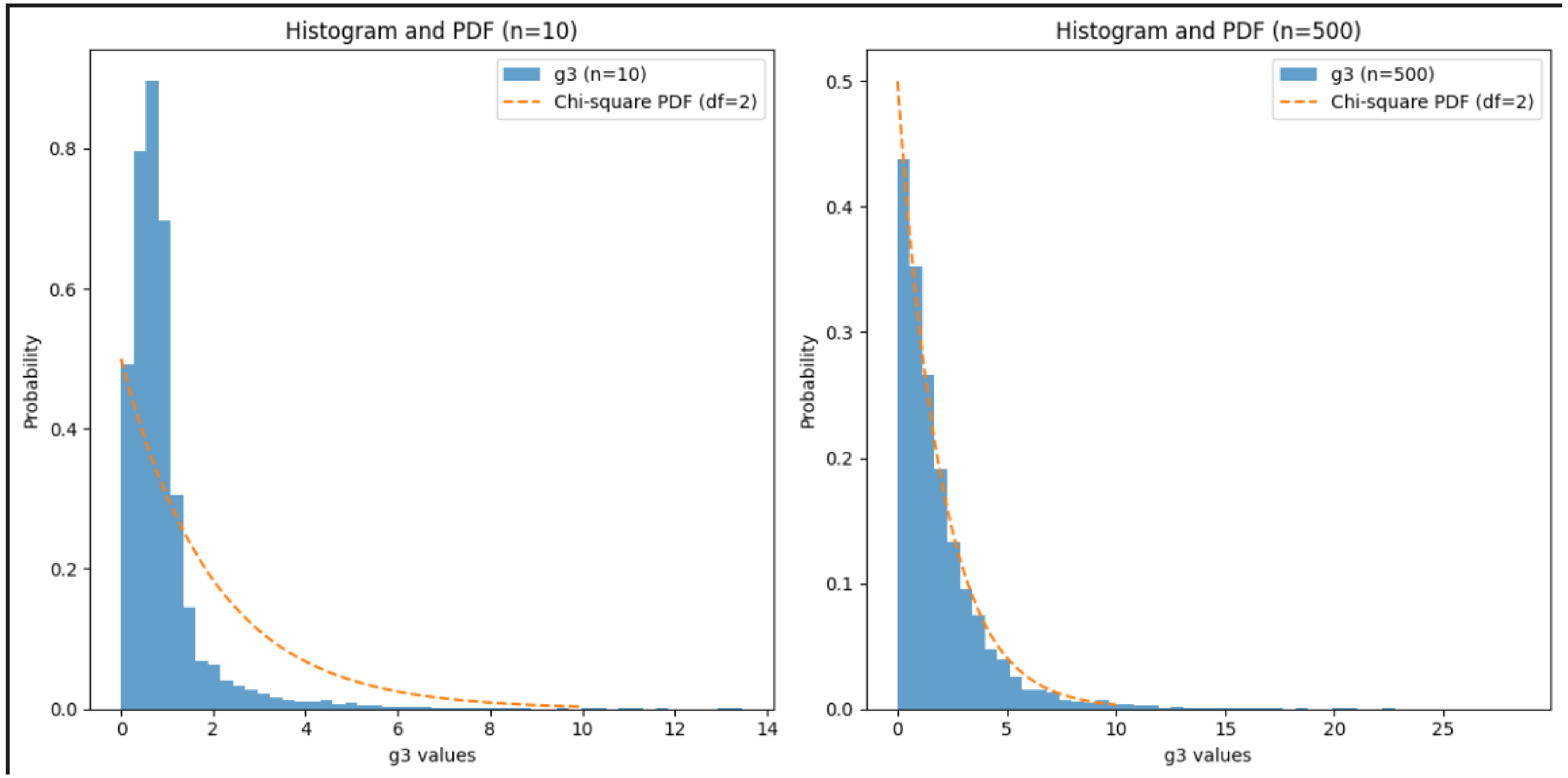
觀察檢定統計量G3的直方圖與對應的卡方PDF



觀察檢定統計量 G3 的直方圖和對應的卡方 PDF

樣本數 n=10 時統計量G3的直方圖不太接近卡方PDF

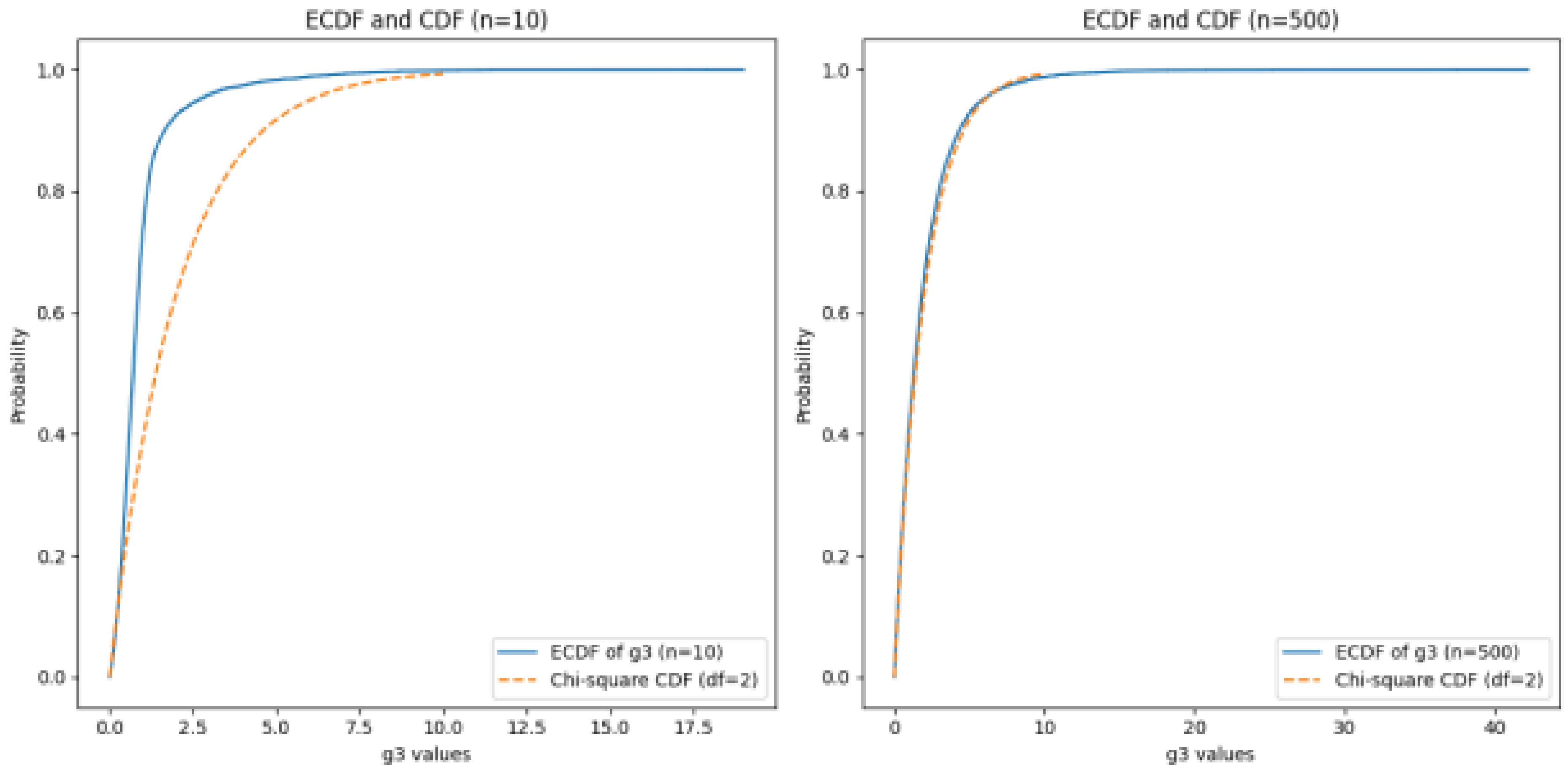
當樣本數增大到 500 時，比在樣本數10時更接近



觀察檢定統計量 G3 的 ECDF 圖和對應的卡方 CDF

統計量的 ECDF 圖在樣本數 n = 10 時，與卡方 CDF 圖在上升的階段差異明顯

樣本數增大到 500 時，非常接近於卡方的 CDF 圖



改寫驗證程式成
JB_test 函數(x)
副程式

JB_test (x)函數程式碼說明:

1.G3=G1^2+G2^2

2.JB_test(X) 輸入要檢定是否為常態的一組資 料，則可以得出兩個結果，stats 為 檢定統計量的值，p-value 為檢定的 p-value

3.參數:

- 樣本數 n = 1000
- 模擬次數N=50,000
- alpha=0.05

➡ 之後只需要引用此副程式可以簡化重複的程式碼，提升執行效率

```python
def JB_test(X):
    n = X.shape[0]
    G1 = np.sqrt(n / 6) * skew(X, bias=False)
    G2 = np.sqrt(n / 24) * kurtosis(X, bias=False)
    stats = G1 ** 2 + G2 ** 2
    p_value = 1 - chi2.cdf(stats, df=2)
    return stats, p_value
# 輸入資料
n = 50 # sample size
N = 50000 # number of samples
X = norm.rvs(size=(N, n))
stats, p_value = JB_test(X)
print(f"Stats: {stats} ")
print( f"p-value: {p_value}")
```

```
Stats: [1.55984365e+00 7.48462405e-01 3.00286022e+00 3.57683716e+00
 1.90983816e+00 3.01365776e-01 4.15903784e-03 6.83494056e-01
 2.90802314e+00 1.37902134e+00 7.02992463e+00 7.88876322e-01
 2.23288553e+00 1.85857150e+00 1.62939283e+00 9.84292773e-01
p-value: [0.45844185 0.68781787 0.22281129 0.16722441 0.38484328 0.860120
 0.99792264 0.71052793 0.23363118 0.50182157 0.02974892 0.67405865
 0.32744252 0.39483562 0.44277373 0.61131287 0.91233609 0.45548292
 0.58946067 0.56784078 0.74400951 0.49960852 0.78566858 0.18554943
```

第二部分:
以蒙地卡羅實驗驗證
J-B 檢定統計量

檢驗檢定統計量

G3 的檢定力

$Power = P(Reject\ H_0 \mid H_a)$ H_0: 資料來自常態 及 H_a: 資料來自其他分配

- 檢定力(Power)為在資料來自其他分配的情況下，拒絕資料來自常態的機率，換言之是該檢定統計量在Ha為真的情況下作出拒絕H0正確決策的機率

- 當 Ha 來自常態時（也就是資料來自 H0 的意思），此時的 Power 又稱為顯著水準，且理論上， Power 應該維持在所設定的型一誤 alpha，即0.05

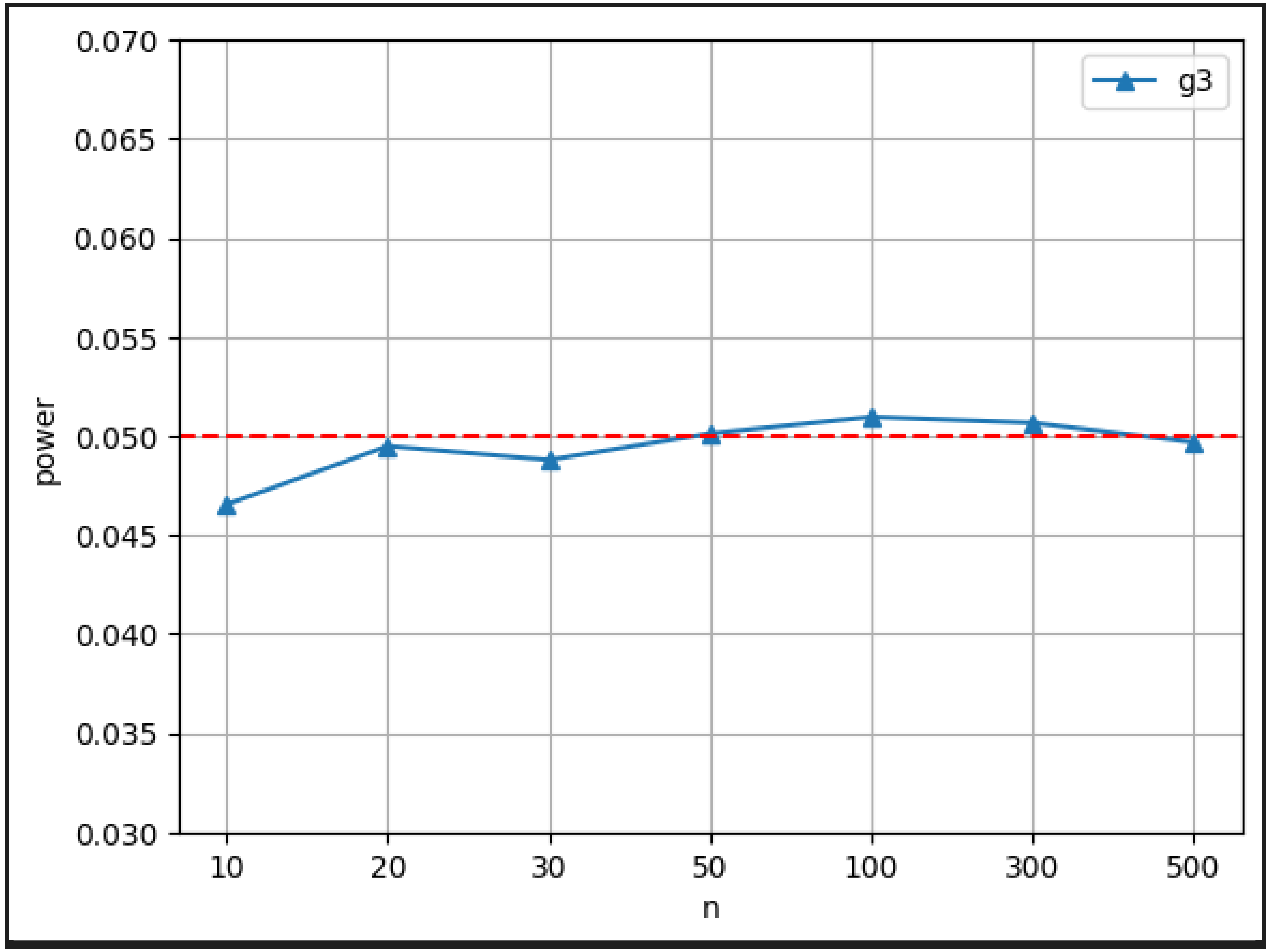
- 檢定統計量是根據 H0 為真的條件下得到的，因此檢定統計量的先要條件是維持既定的顯著水準，這是做後續檢定力的大前提

檢驗檢定統計量 G3 的檢定力。採蒙地卡羅模擬方式，步驟如下：

- 從下列的分配母體中抽樣：N(0,1), T(3), T(10), T(30), U(0,1), chi(8)
- 抽樣數 n=10, 20, 30, 50, 100, 300, 500。
- 實驗次數 N = 50000
- 型一誤 alpha = 0.05
- 以下分成兩類，一類是當資料來自 H0，另一類是當資料來自 Ha

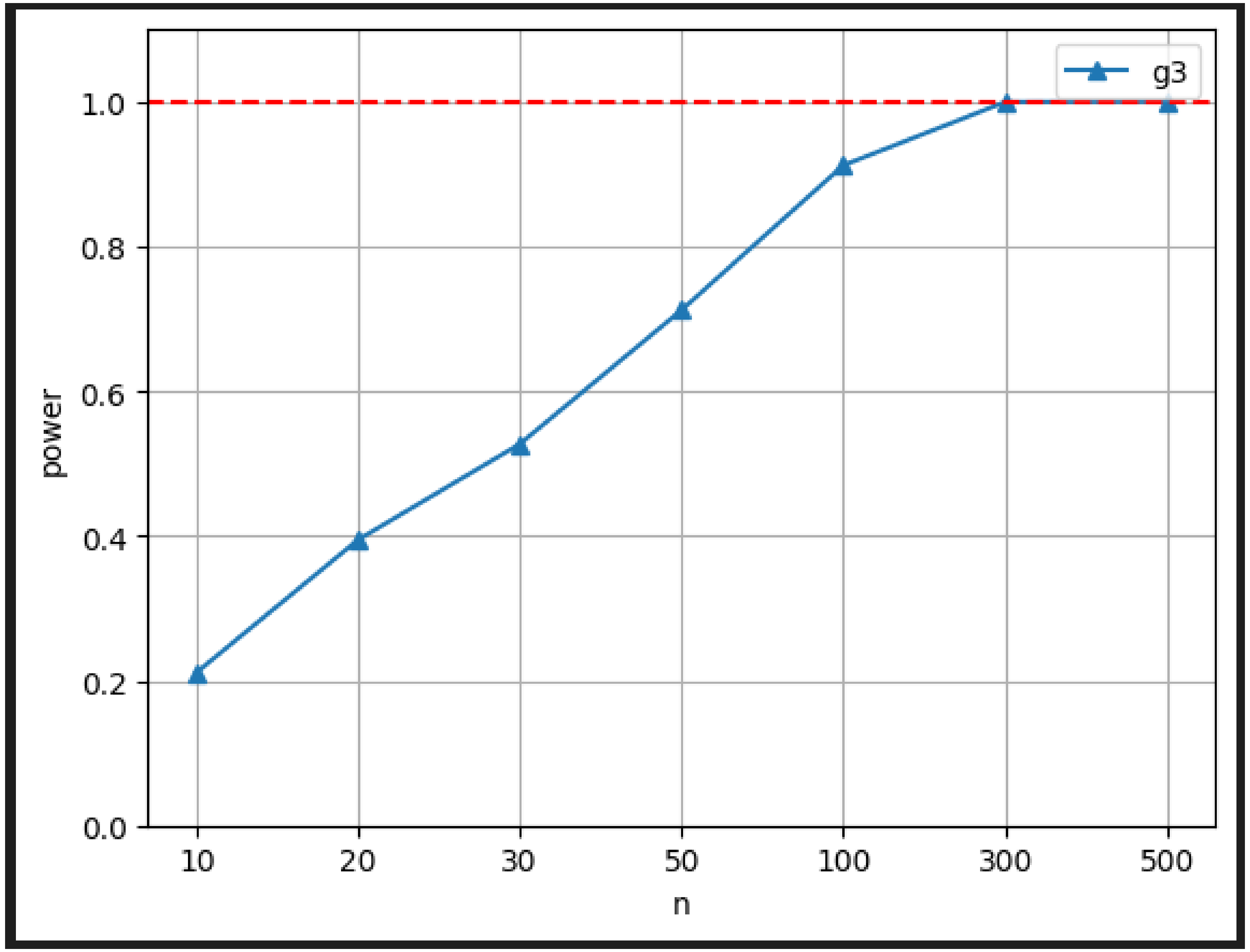
當資料來自 H0，檢定力也稱「顯著水準」，理論上必須維持在 alpha值

母體：N(0,1)



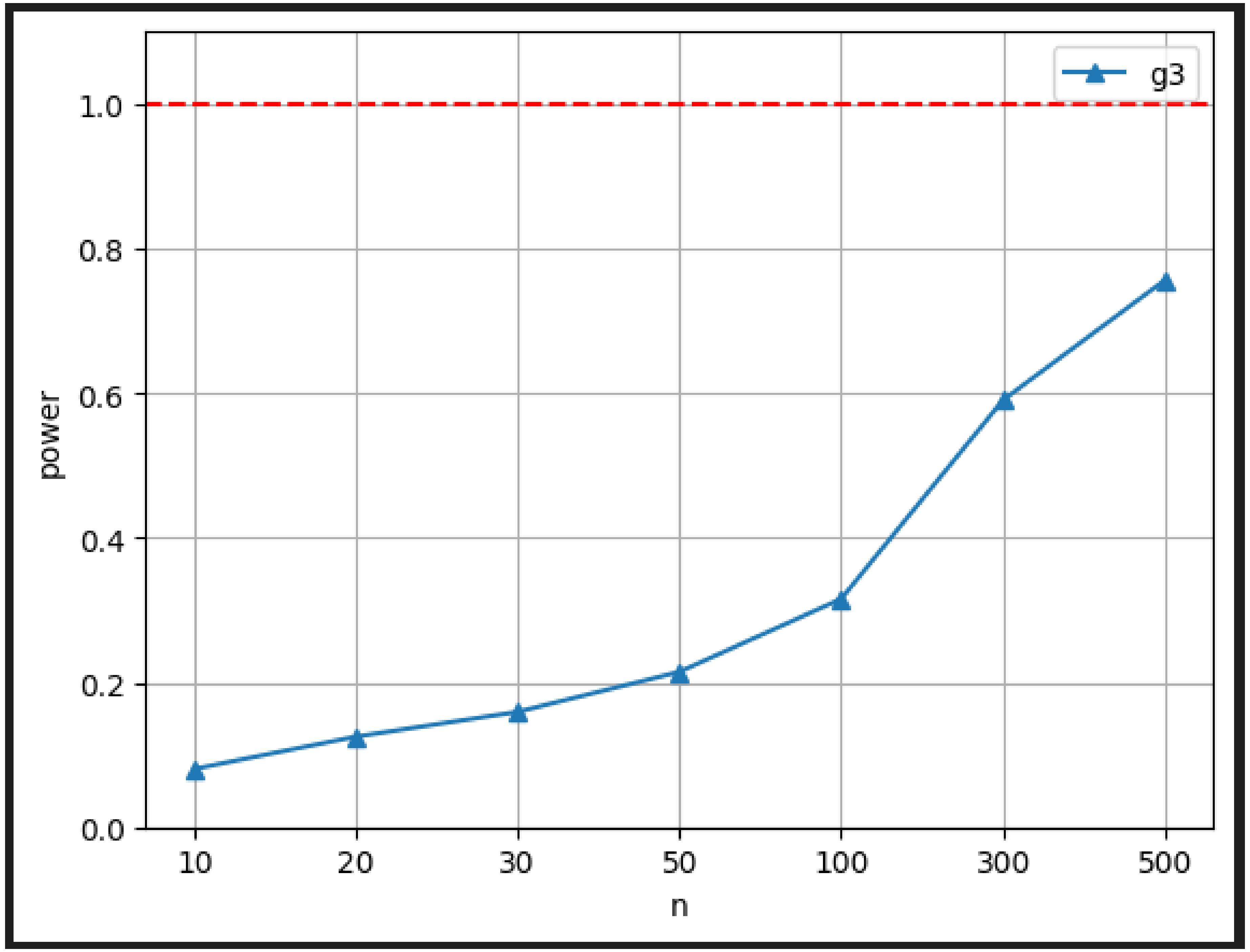
當資料來自 Ha，檢定力隨著樣本數增加而變大

母體：T(3)



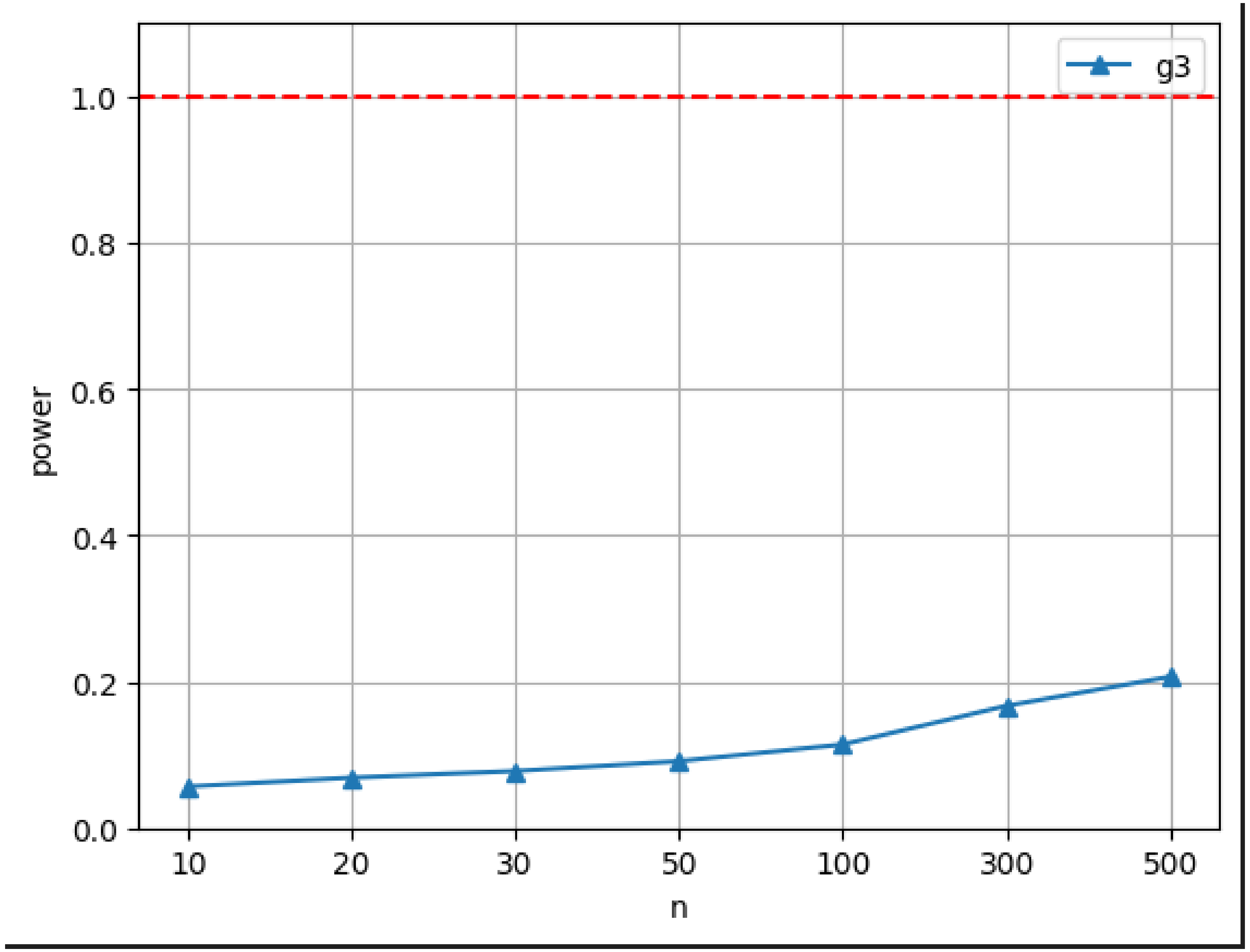
當資料來自 Ha，檢定力隨著樣本數增加而變大

母體：T(10)



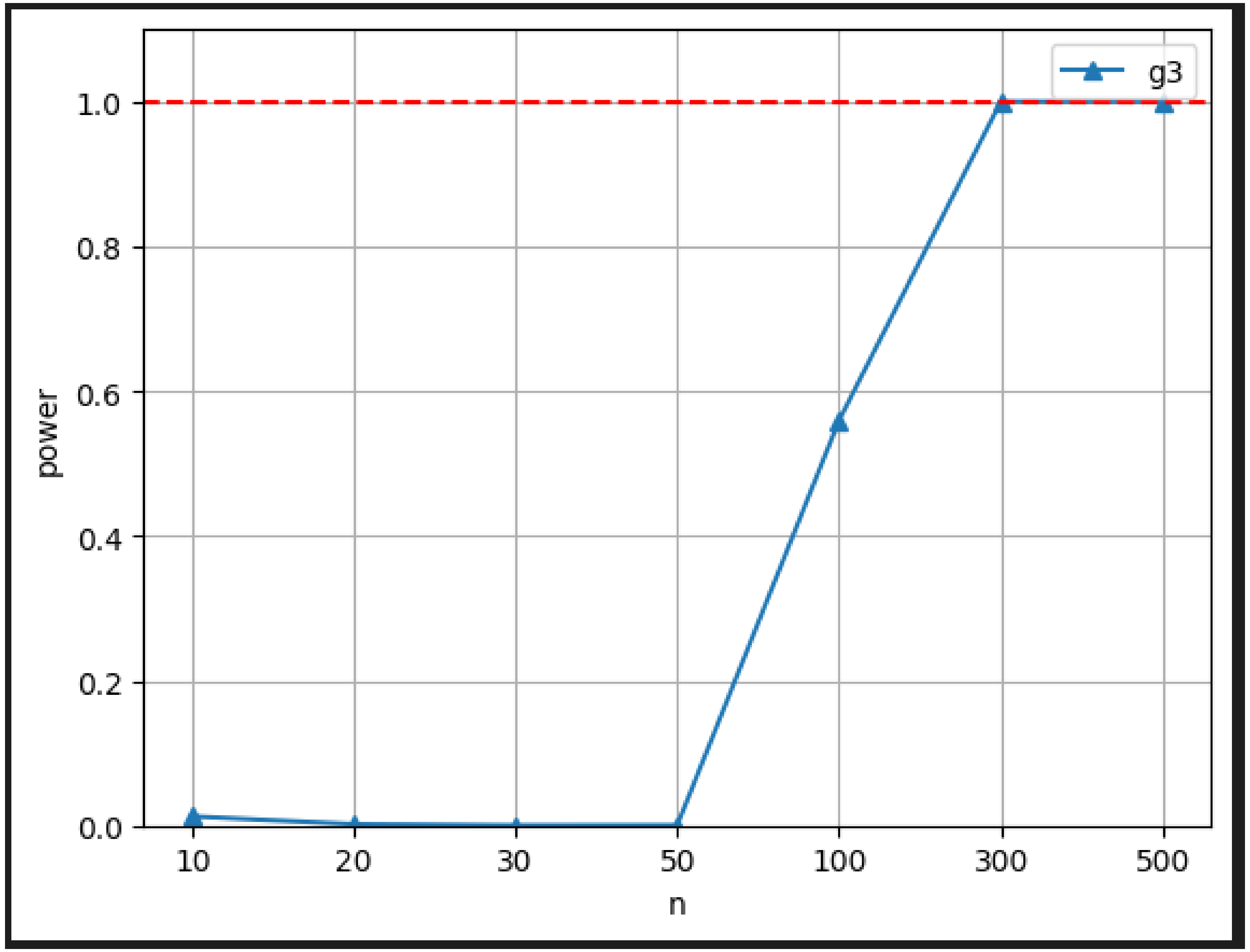
當資料來自 Ha，檢定力隨著樣本數增加而變大

母體：T(30)



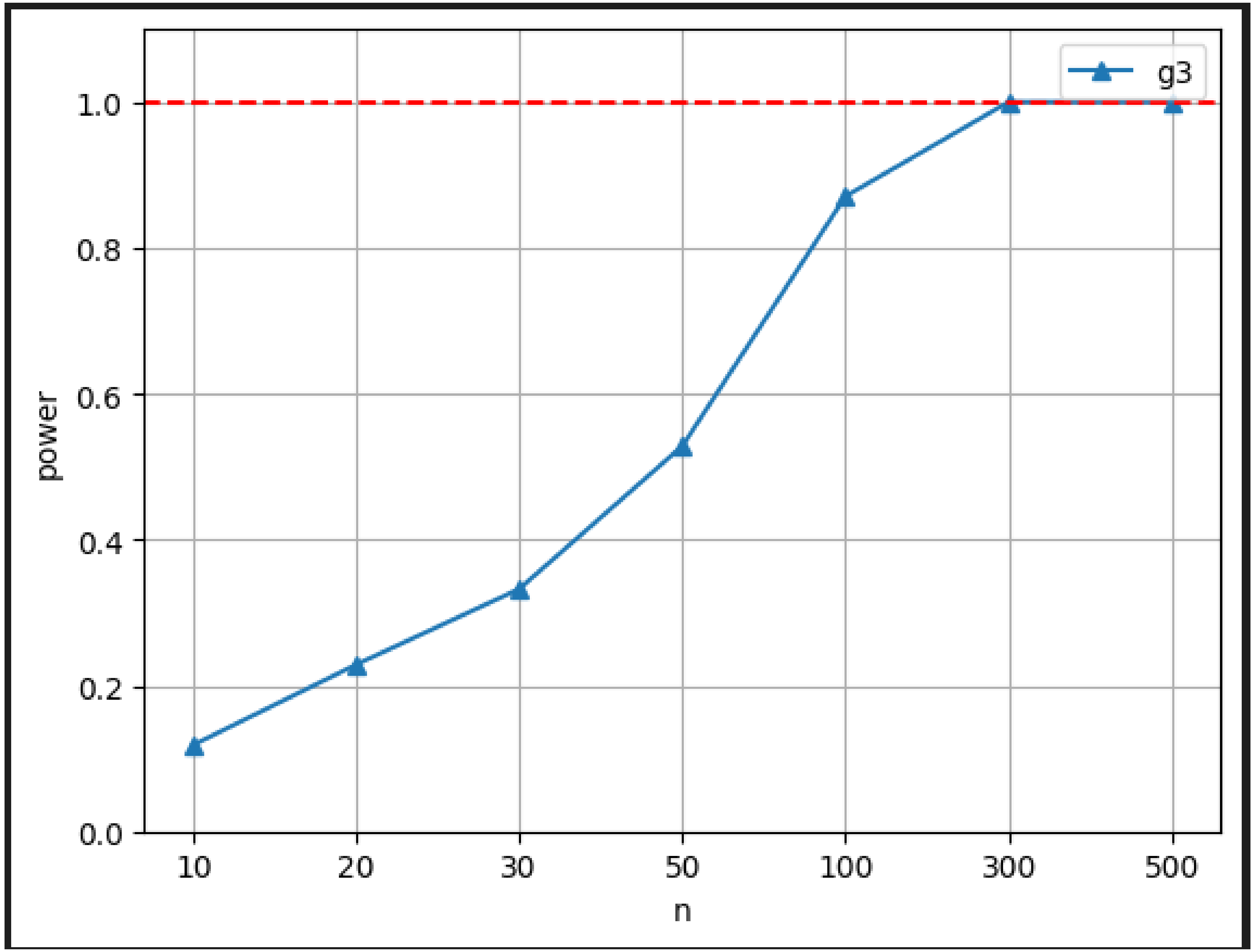
當資料來自 Ha，檢定力隨著樣本數增加而變大

母體：U(0, 1)



當資料來自 Ha，檢定力隨著樣本數增加而變大

母體：chi^2(8)



比較 G1 v.s. G2 v.s. G3 在各種樣本數下從不同母體抽樣的檢定力

定義一個名為 plot_reject_ratio 的函數，該函數用於繪製在不同樣本大小下，對三個統計量（G1、G2、和 G3 統計量）進行的檢定的結果。

函數參數：

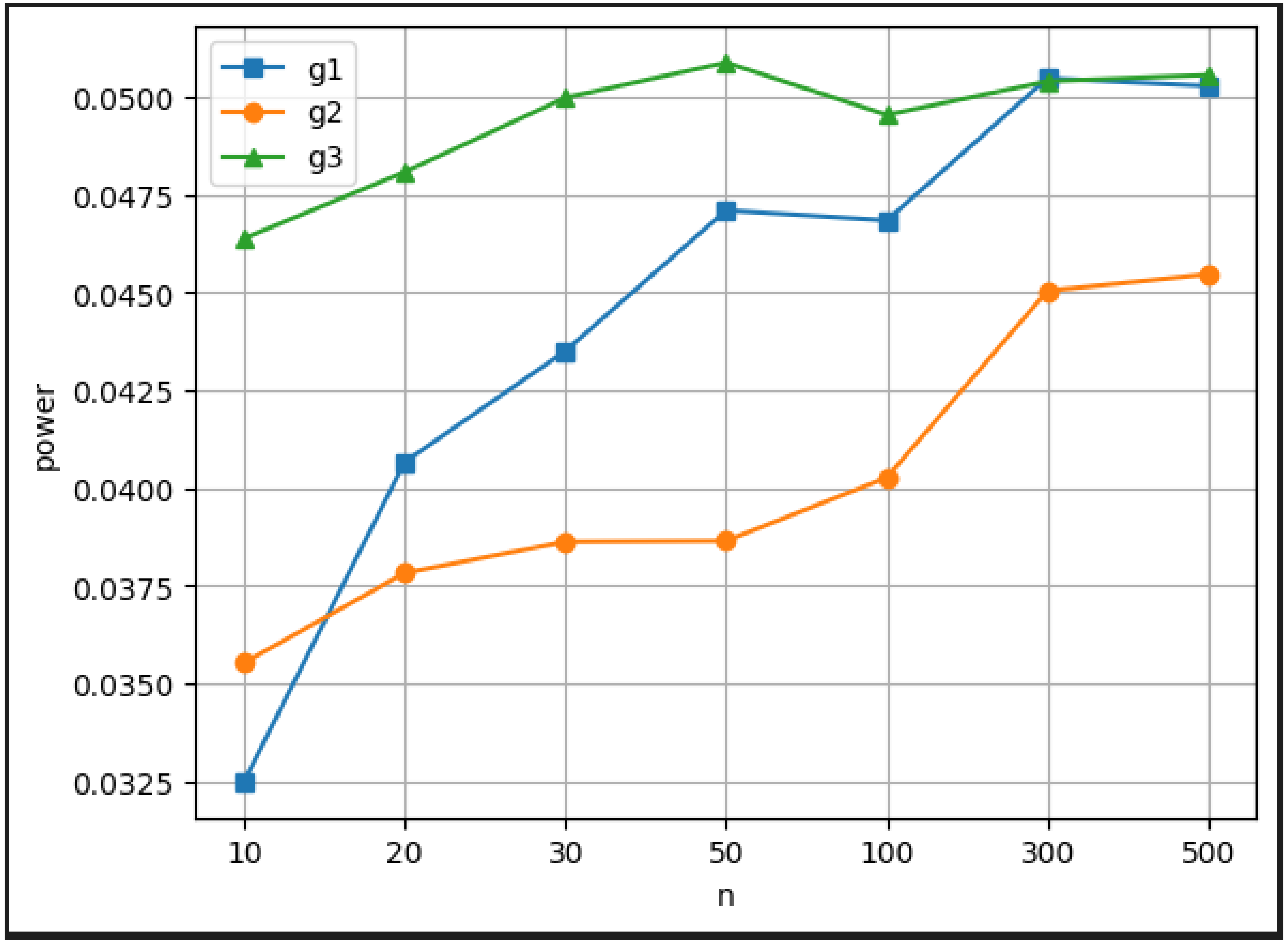
1. **n_values[10, 20, 30, 50, 100, 300, 500]: 一個包含不同樣本大小的列表。**

2. **N（50000）：每個樣本大小的樣本數。**

3. **alpha（0.05）：顯著性水平。**

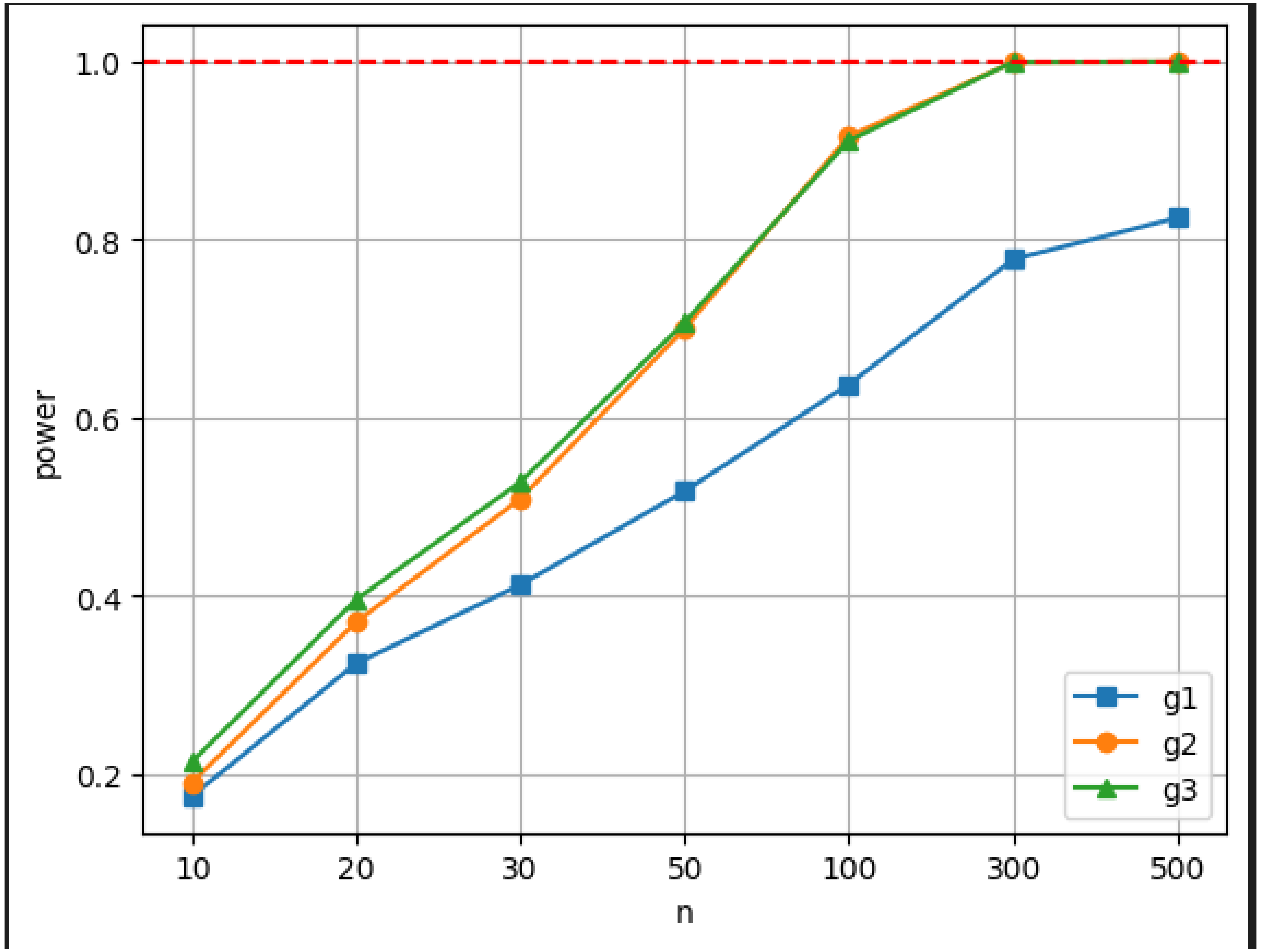
比較三種檢定統計量從不同母體抽樣下的檢定力

母體：N(0, 1)



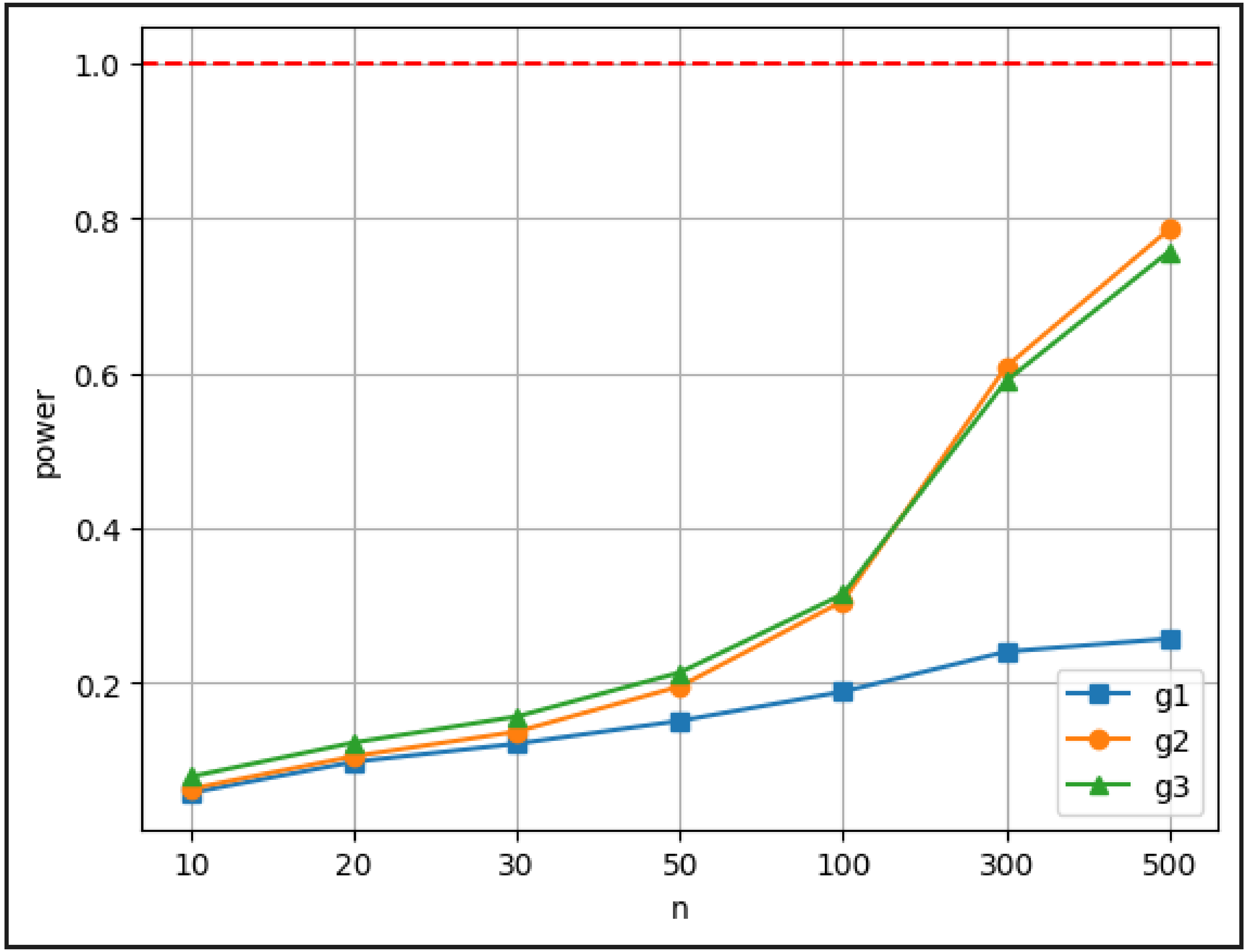
比較三種檢定統計量從不同母體抽樣下的檢定力

母體：T(3)



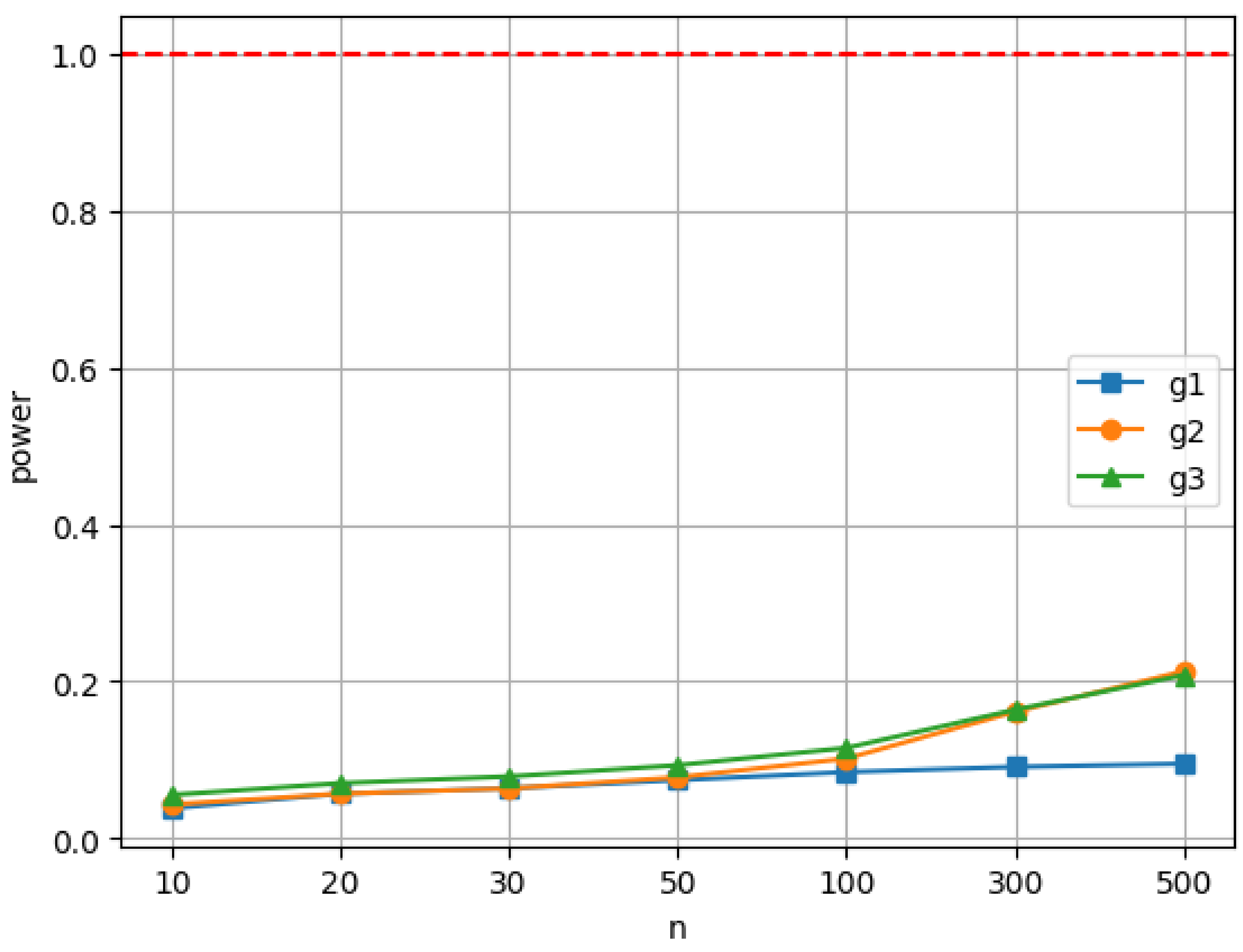
比較三種檢定統計量從不同母體抽樣下的檢定力

母體：T(10)



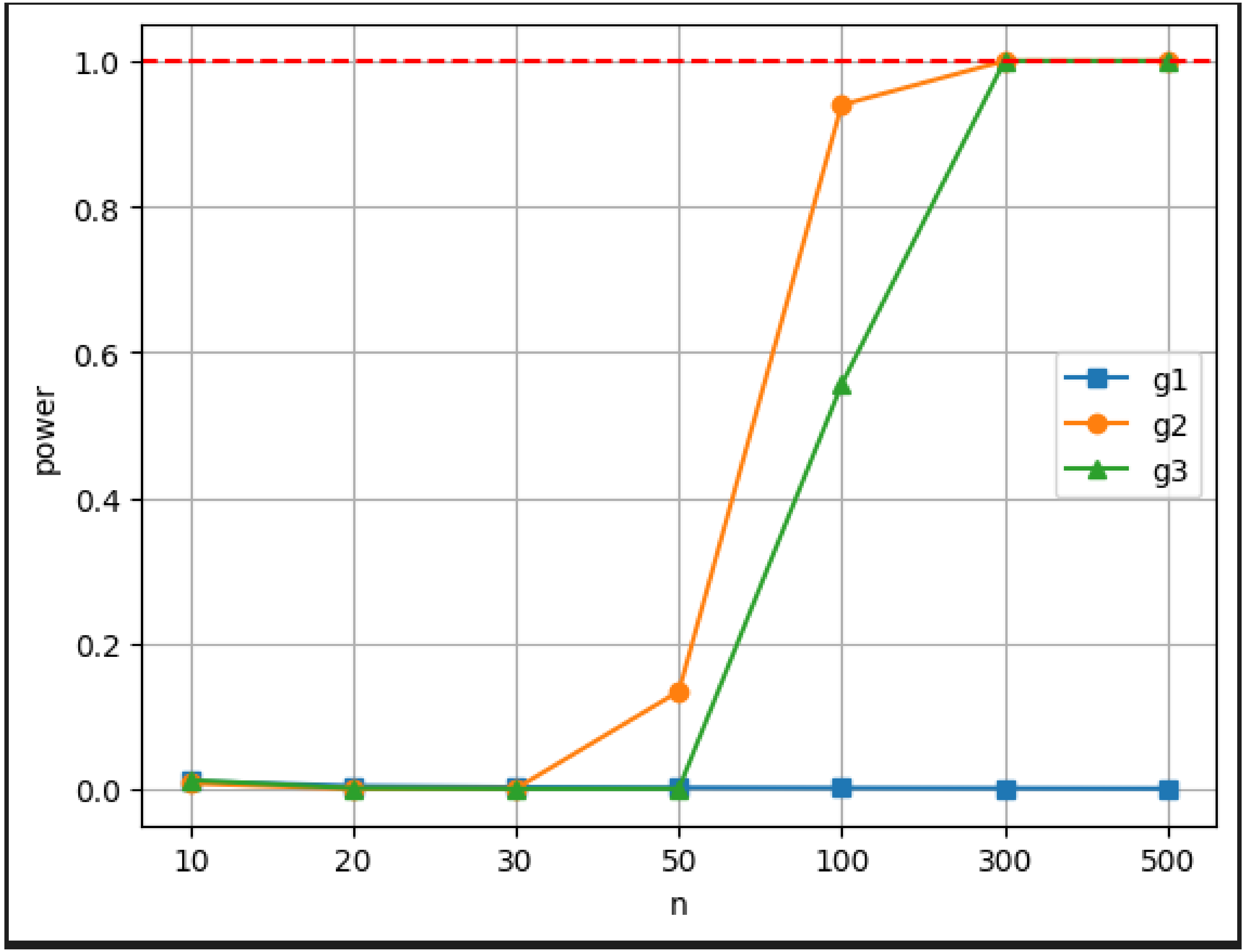
比較三種檢定統計量從不同母體抽樣下的檢定力

母體：T(30)



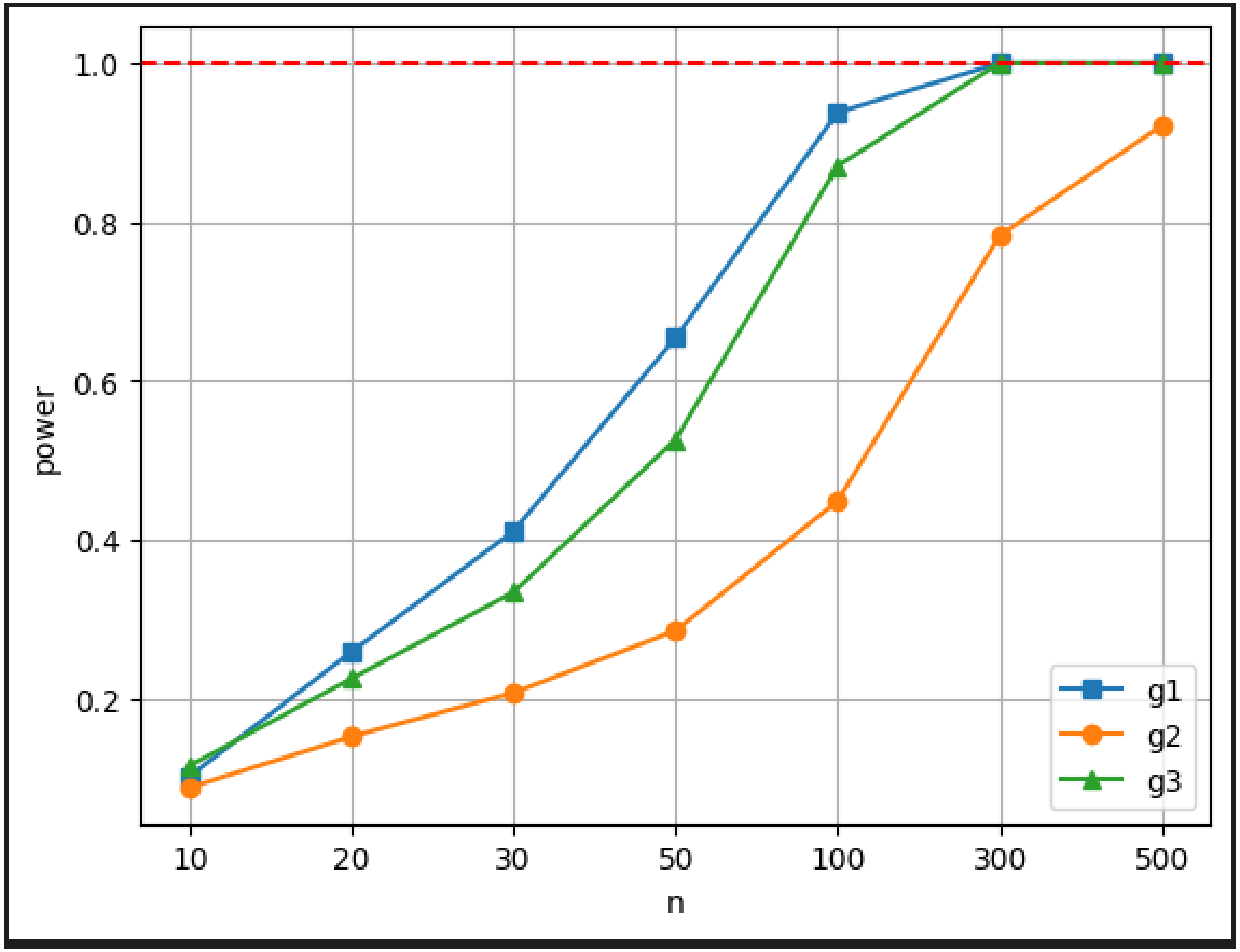
比較三種檢定統計量從不同母體抽樣下的檢定力

母體：U(0, 1)



比較三種檢定統計量從不同母體抽樣下的檢定力

母體：chi^2(8)



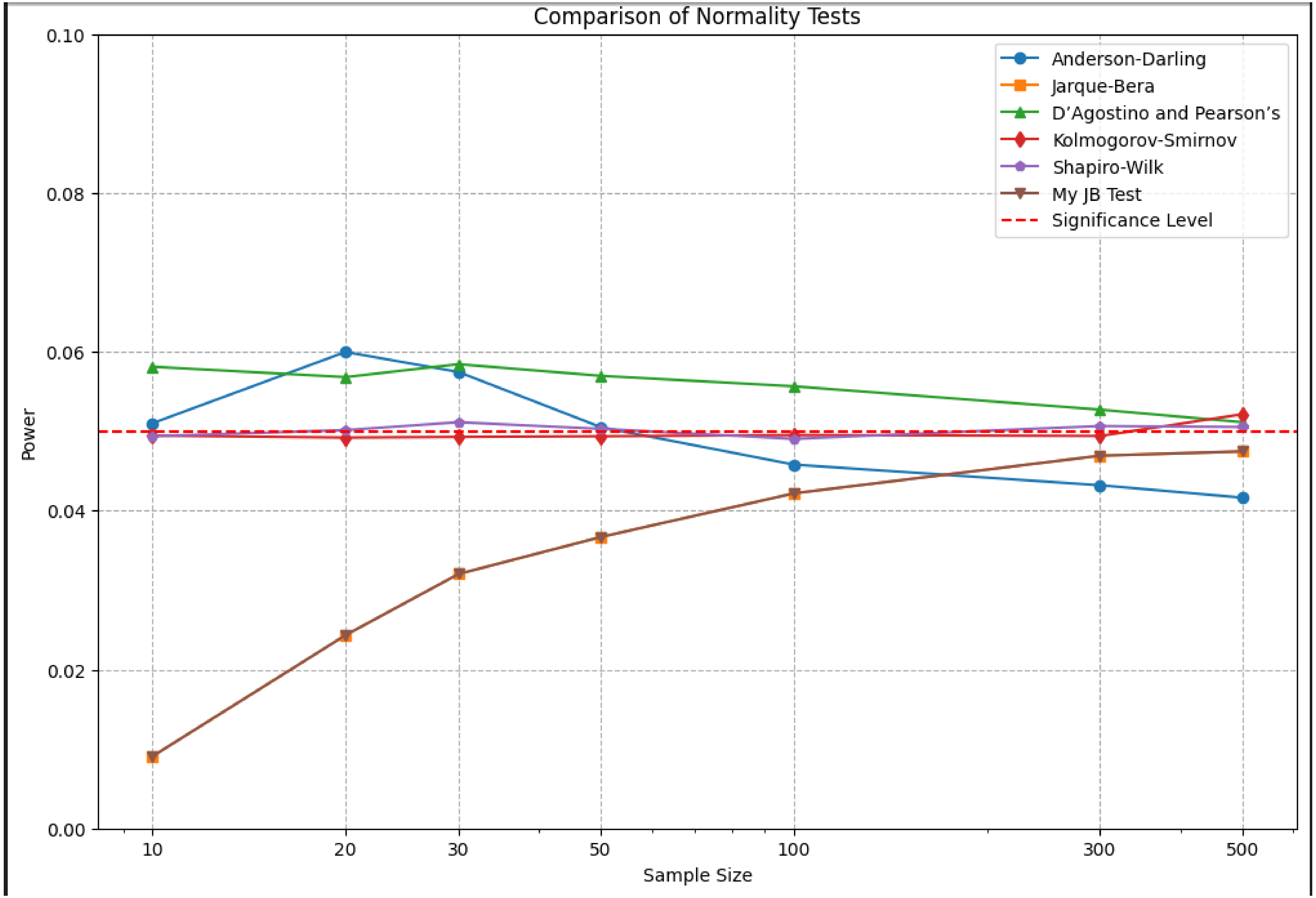
JB test vs常態檢定的產品

程式碼說明:

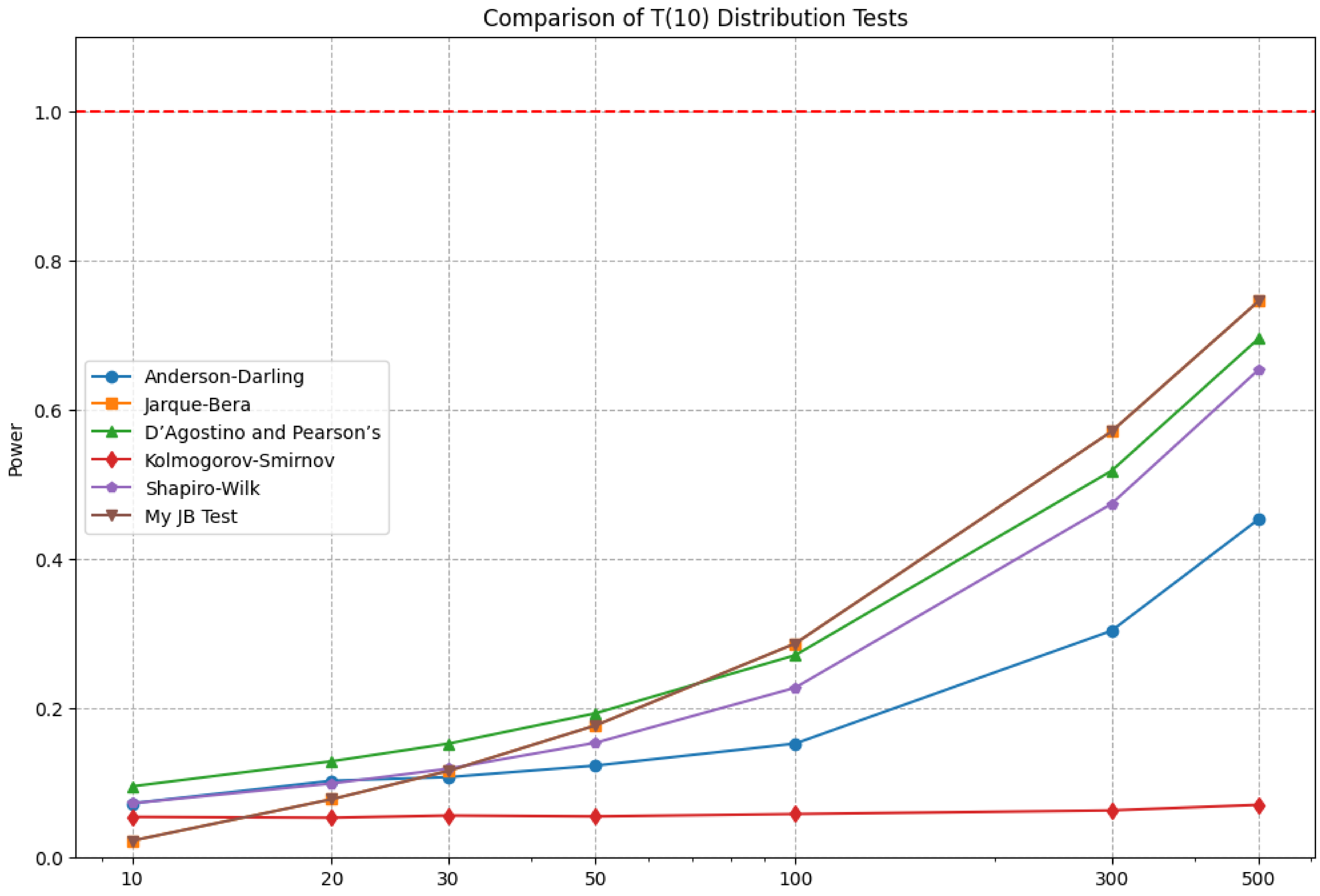
1. **常態性檢定方法: Anderson-Darlin, Jarque-Bera, D’Agostino and Pearson’s, Kolmogorov-Smirnov, Shapiro-Wilk (from scipy.stats)**
2. **模擬了不同樣本大小下的檢定力**
3. **參數:**

- **樣本數 n = 10, 20, 30, 50, 100, 300, 500, 1000。**
- **針對每個樣本數 n，模擬次數皆為 N=50,000。**
- **alpha=0.05**

母體：N(0, 1)


Comparison of Normality Tests
0.10
0.08
0.06
0.04
0.02
0.00
Power
10
20
30
50
100
300
500
Sample Size
Anderson-Darling
Jarque-Bera
D'Agostino and Pearson's
Kolmogorov-Smirnov
Shapiro-Wilk
My JB Test
Significance Level

母體：T(10)


Comparison of T(10) Distribution Tests
Power
1.0
0.8
0.6
0.4
0.2
0.0
10
20
30
50
100
300
500
Anderson-Darling
Jarque-Bera
D'Agostino and Pearson's
Kolmogorov-Smirnov
Shapiro-Wilk
My JB Test

總結

第一部分觀察 G1, G2, G3 的抽樣分配與樣本的關係

第二部分則是是以蒙地卡羅實驗驗證 J-B 檢定統計量

謝謝!