Non-Parametric Tests for Scale Families

Dishari Datta, Dyuti Manik, Iman Kalyan Dutta

Indian Statistical Institute

April 16, 2025

Outline

1. Introduction to the Problem
2. Klotz Statistic
3. Mood Statistic
4. Savage Statistic
5. Asymptotic Size
6. Asymptotic Power
7. Innovation

Problem Setup

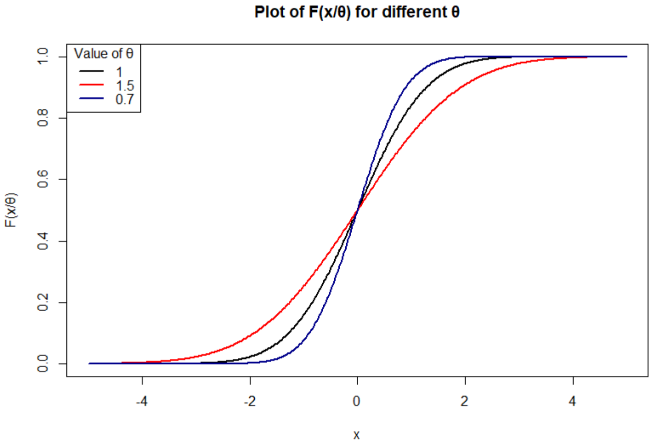
We consider two independent samples:

$$X_1, X_2, \ldots, X_n \sim F(x)$$
$$\mathrm{Y}_1, Y_2, \ldots, Y_m \sim G(x), \quad \text{where } G(x) = F\left(\frac{x}{\theta}\right)$$

Our objective is to test the following hypotheses:

$$H_0 : \theta = 1 \qquad \text{against} \qquad \begin{aligned} H_1 &: \theta > 1 \\ H_2 &: \theta < 1 \\ H_3 &: \theta \neq 1 \end{aligned}$$

Visualising F(x/θ)



When $\theta > 1$, the sample Y is more dispersed than X.
When $\theta < 1$, the sample X is more dispersed than Y.

Alternatives chosen and Tests used

For the purpose of this study, we focus on the one-sided alternative:

$$H_1 : \theta > 1$$

In order to perform the test H_0 vs H_1, we use the following statistics:

– Klotz Statistic

– Mood Statistic

– Savage Statistic

Notations

We rank all the observations from the two samples together, from the smallest to the largest.

Let $R_1, R_2, \ldots, R_n$ be the ranks of $X_1, X_2, \ldots, X_n$.

and $R_{n+1}, R_{n+2}, \ldots, R_{n+m}$ be the ranks of $Y_1, Y_2, \ldots, Y_m$.

in the combined sample.

Let $N = n + m$ denote the total number of observations.

Klotz Statistic

Formula and Distribution under H_0

$$K = \sum_{i=1}^{n} (\Phi^{-1}\left(\frac{R_i}{N+1}\right))^2$$

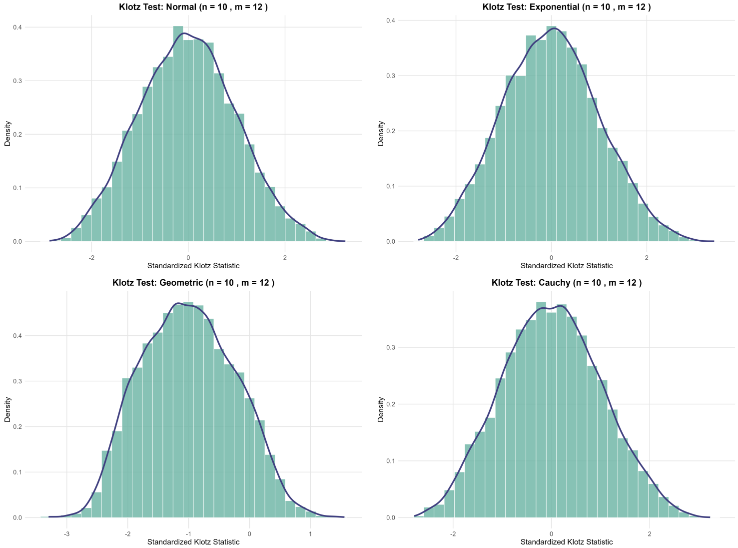
$$\mathbb{E}_{H_0}(K) = \frac{n}{N} \sum_{i=1}^{n} \left[\Phi^{-1}\left(\frac{i}{N+1}\right)\right]^2$$

$$\mathrm{Var}_{H_0}(K) = \frac{nm}{N(N-1)} \sum_{i=1}^{n} \left[\Phi^{-1}\left(\frac{i}{N+1}\right)\right]^4 - \frac{n}{m(N-1)} [\mathbb{E}(K)]^2$$

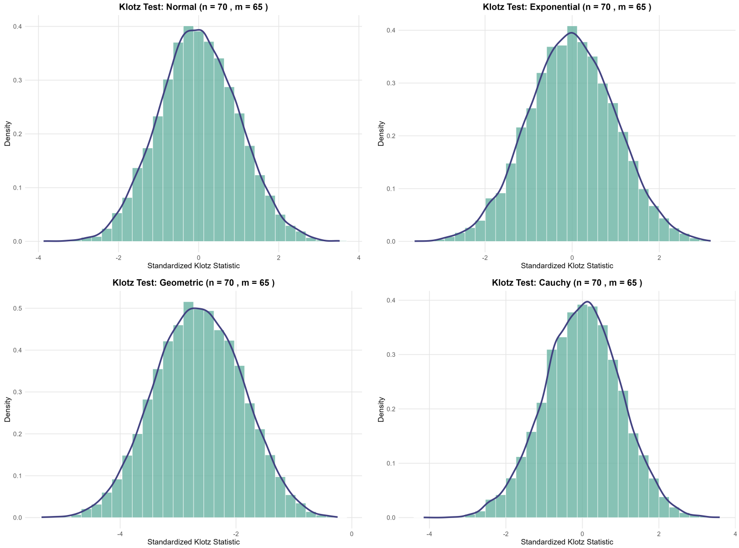
Under H_0, the standardized statistic

$$\frac{K - \mathbb{E}_{H_0}(K)}{\sqrt{\mathrm{Var}_{H_0}(K)}} \sim \mathcal{N}(0,1)$$

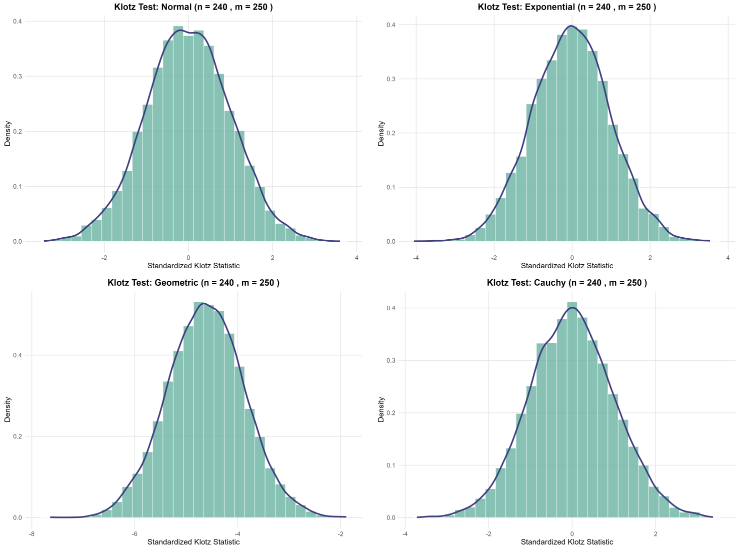
Distribution under H_0 (n = 10, m = 12)



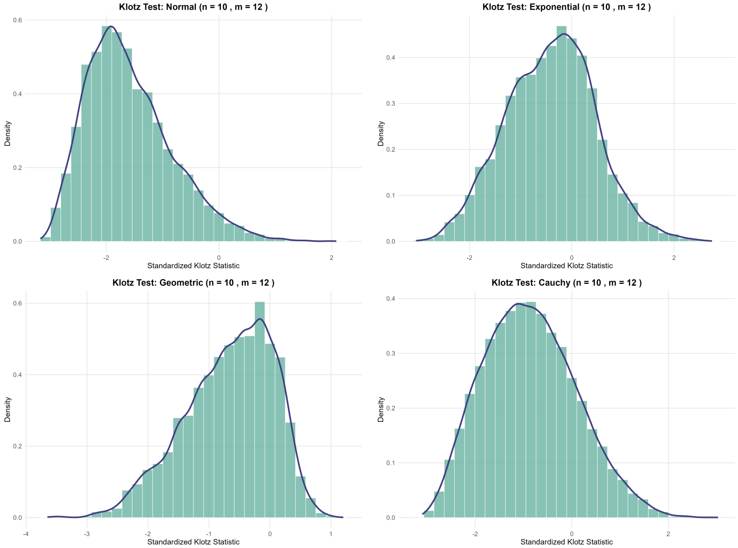
Distribution under H_0 $(n = 70, m = 65)$



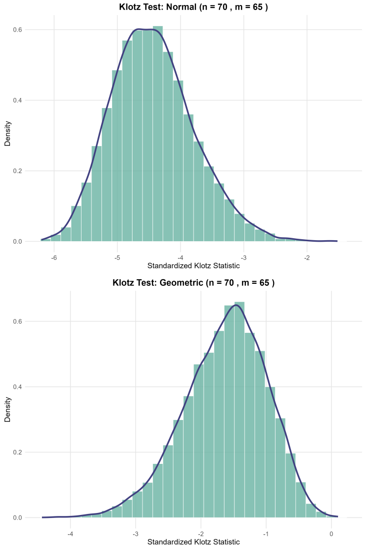
Distribution under H_0 ($n = 240$, $m = 250$)



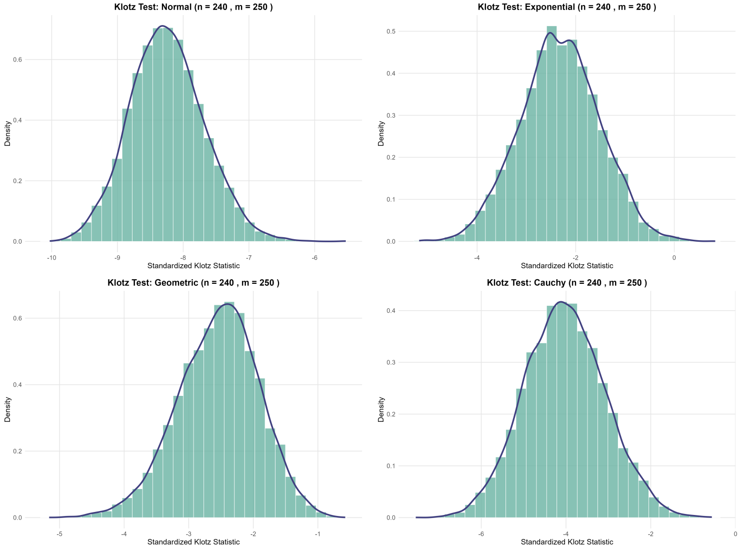
Distribution when $\theta = 2$ (n = 10, m = 12)



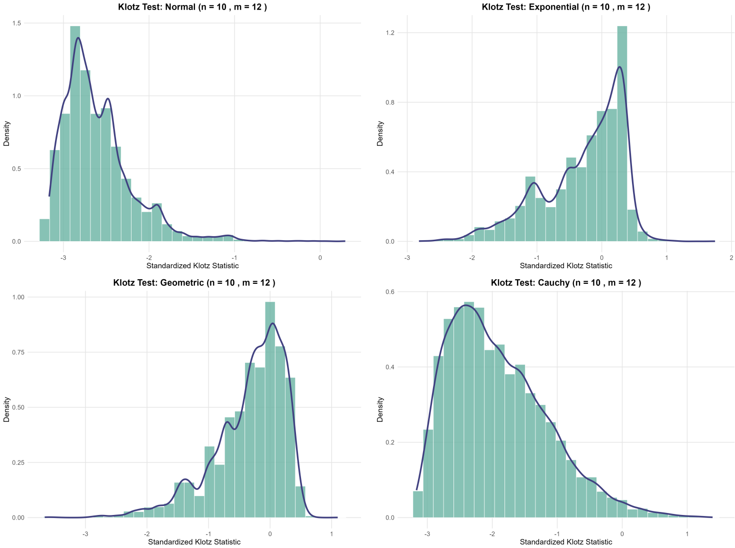
Distribution when $\theta = 2$ (n = 70, m = 65)



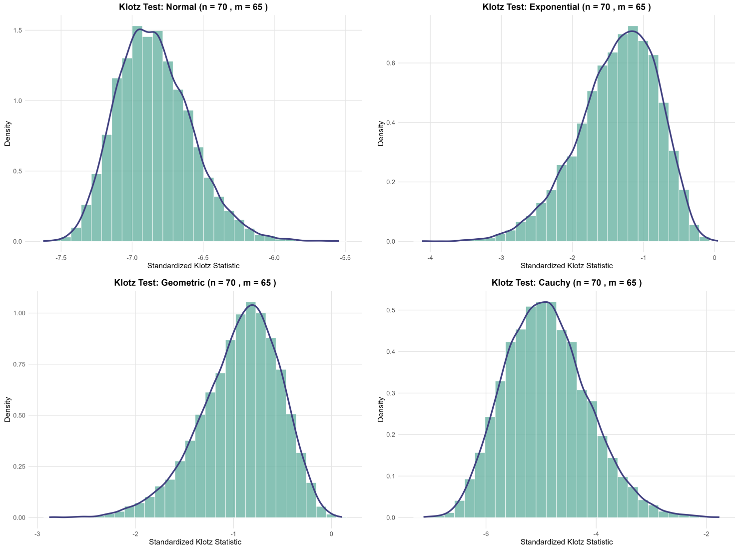
Distribution when $\theta = 2$ (n = 240, m = 250)



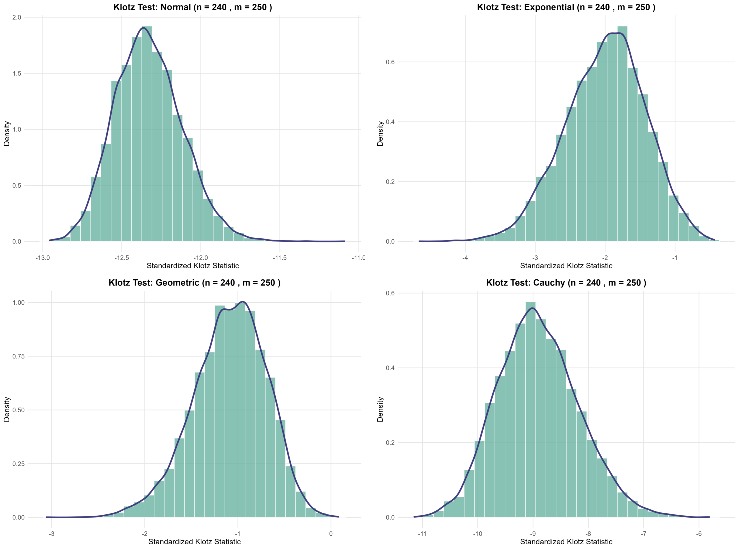
Distribution when $\theta = 6$ (n = 10, m = 12)



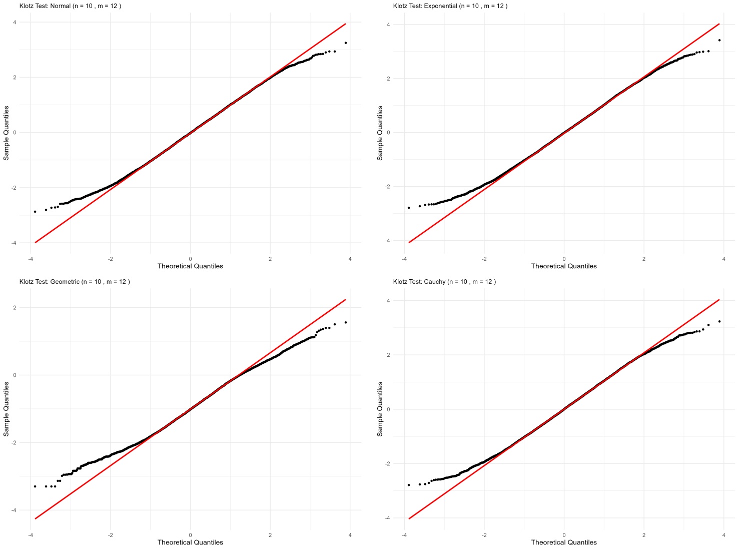
Distribution when $\theta = 6$ (n = 70, m = 65)



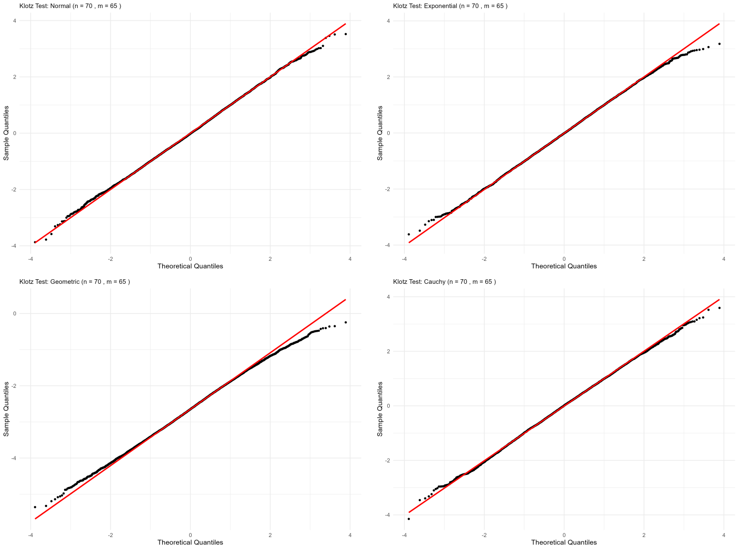
Distribution when $\theta = 6$ (n = 240, m = 250)



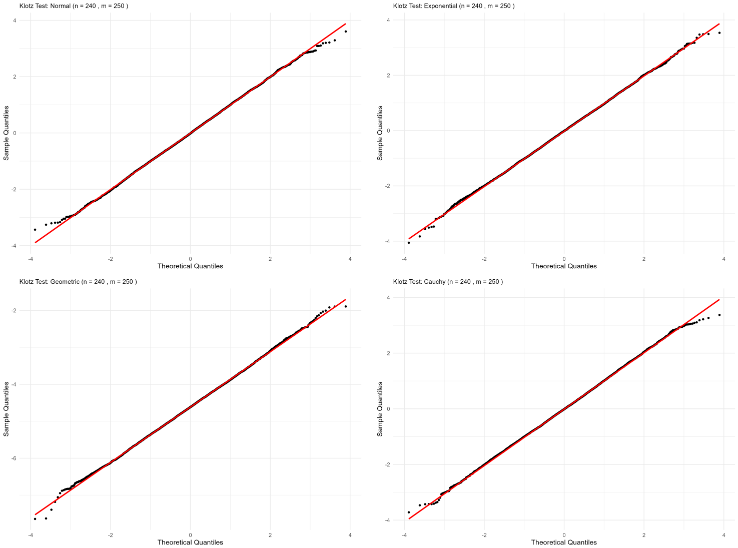
Testing for Normality under H_0 (n = 10, m = 12)



Testing for Normality H_0 ($n = 70$, $m = 65$)



Testing for Normality H_0 (n = 240, m = 250)



Mood Statistic

Formula and Distribution under H_0

$$M = \sum_{i=1}^{n} \left(R_i - \frac{N+1}{2} \right)^2$$

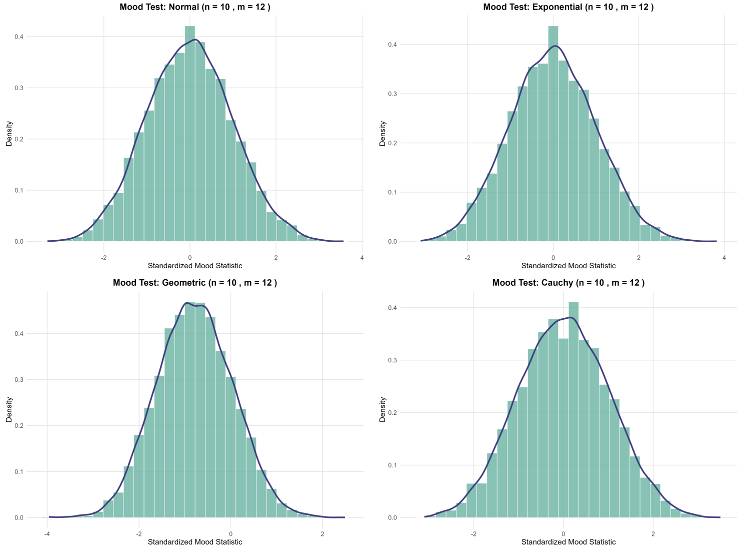
$$\mathbb{E}_{H_0}(M) = \frac{n(N^2-1)}{12}$$

$$\mathrm{Var}_{H_0}(M) = \frac{nm(N+1)(N^2-4)}{180}$$

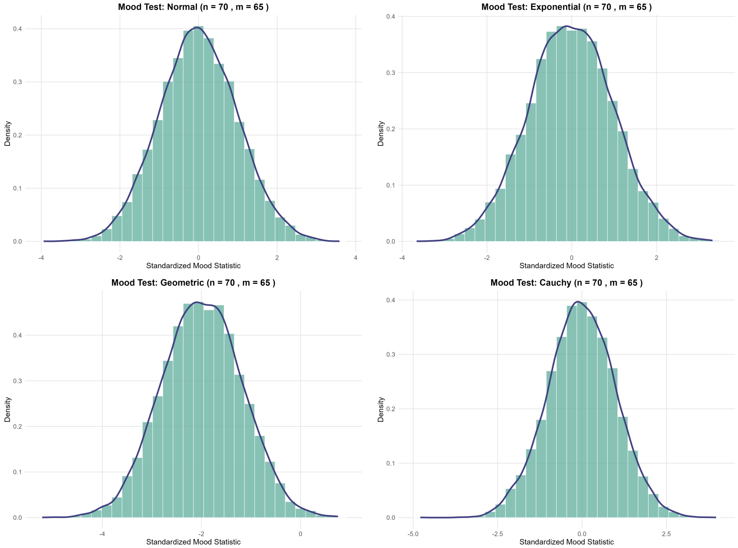
Under H_0, the standardized statistic

$$\frac{M - \mathbb{E}_{H_0}(M)}{\sqrt{\mathrm{Var}_{H_0}(M)}} \sim \mathcal{N}(0,1)$$

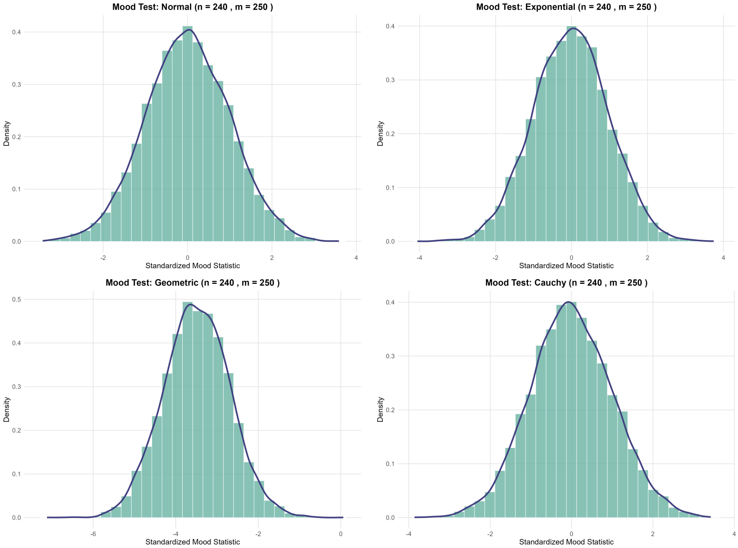
Distribution under H_0 ($n = 10$, $m = 12$)



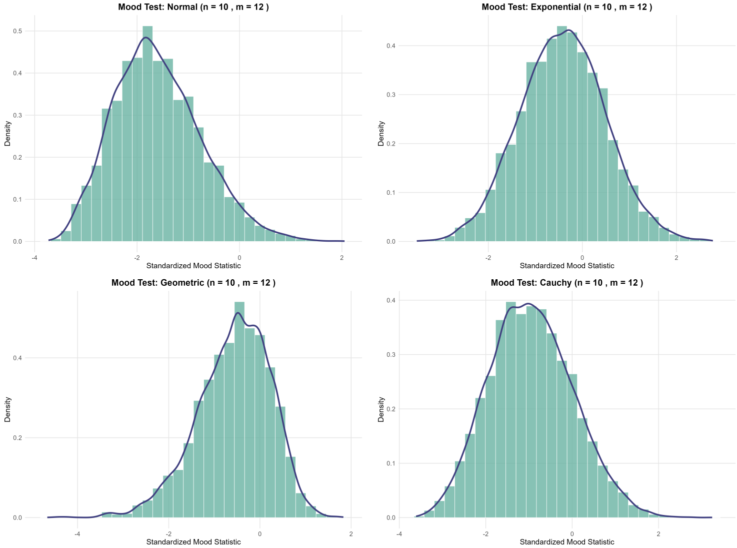
Distribution under H_0 ($n = 70$, $m = 65$)



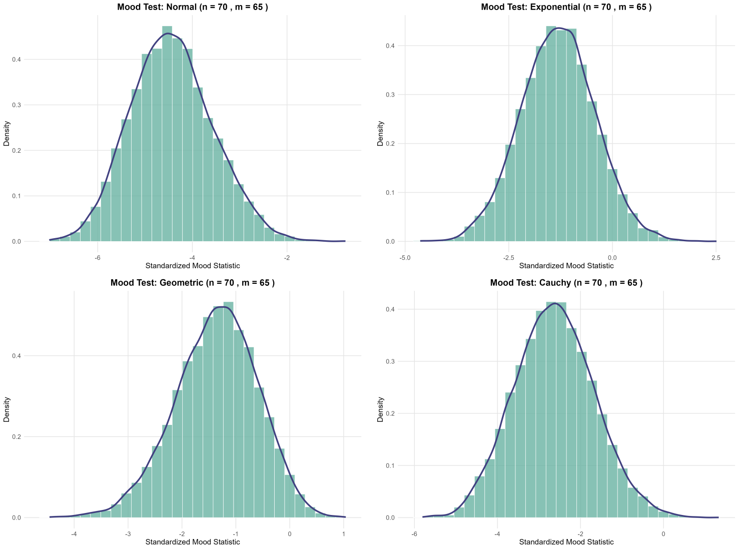
Distribution under H_0 $(n = 240, m = 250)$



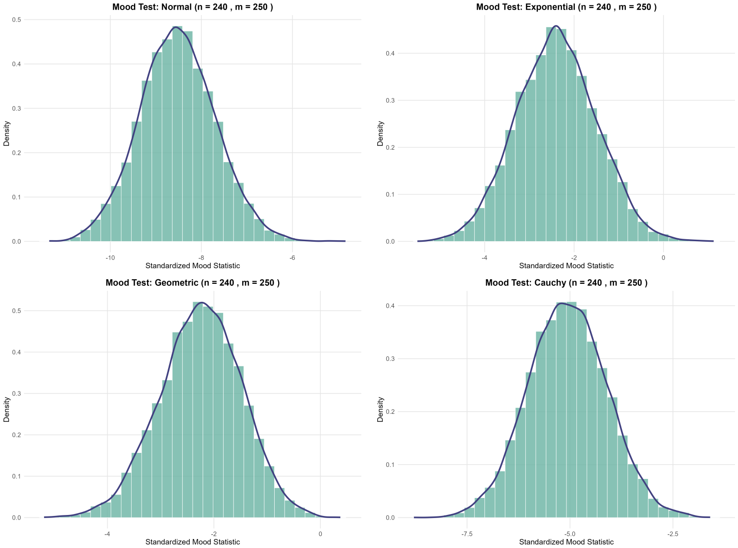
Distribution when $\theta = 2$ (n = 10, m = 12)



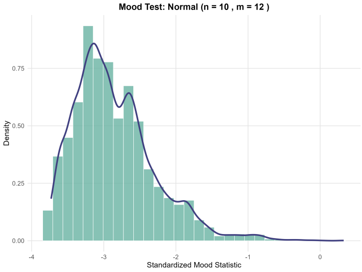
Distribution when $\theta = 2$ (n = 70, m = 65)

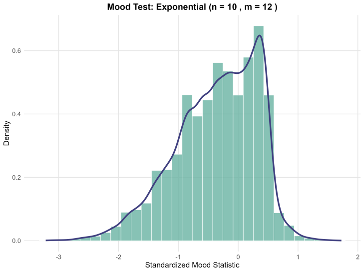


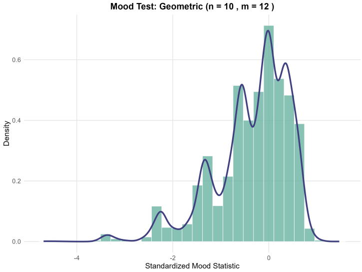
Distribution when $\theta = 2$ (n = 240, m = 250)

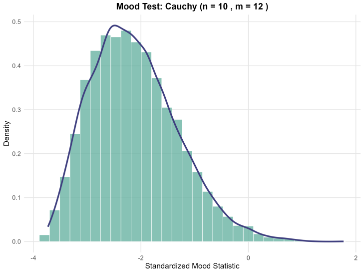


Distribution when $\theta = 6$ (n = 10, m = 12)

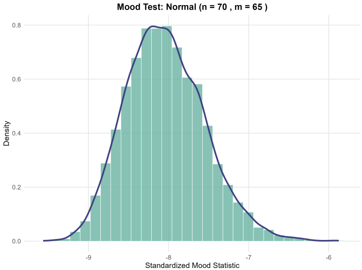


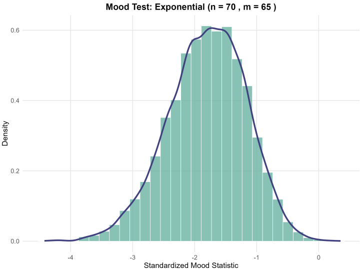


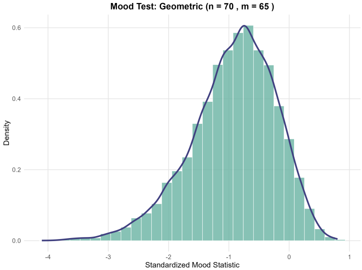


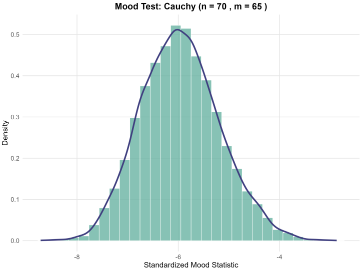


Distribution when $\theta = 6$ (n = 70, m = 65)

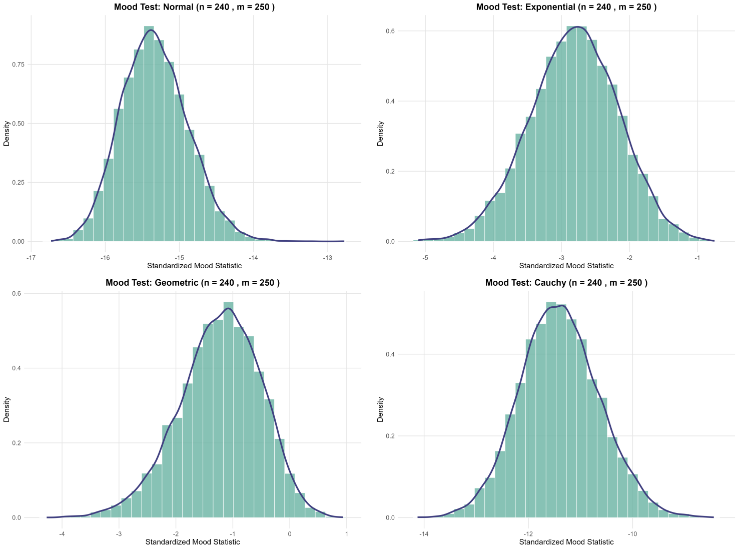




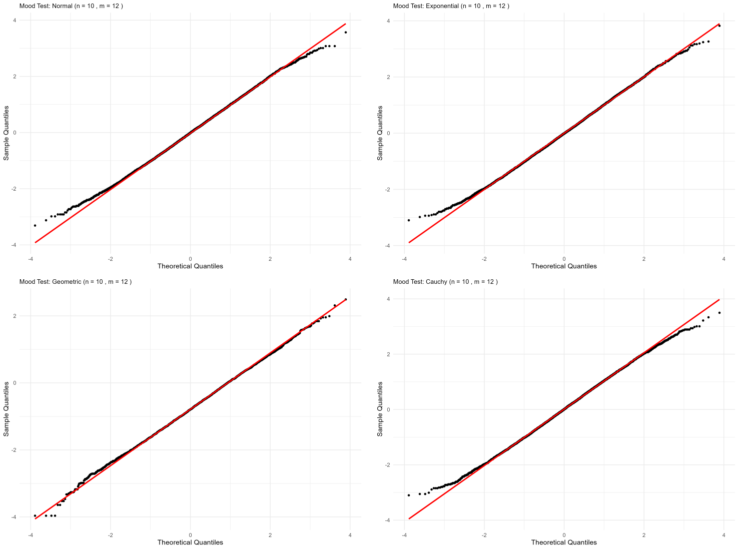




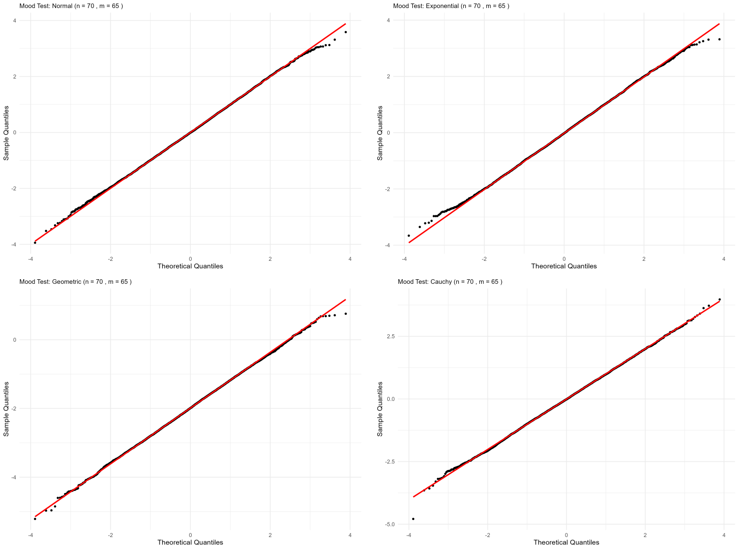
Distribution when $\theta = 6$ (n = 240, m = 250)



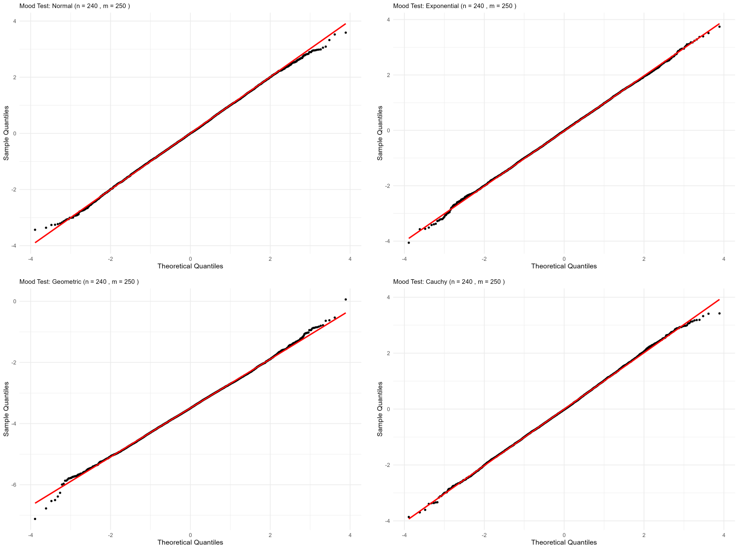
Testing for Normality under H_0 $(n = 10, m = 12)$



Testing for Normality H_0 $(n = 70, m = 65)$



Testing for Normality H_0 ($n = 240$, $m = 250$)



Savage Statistic

Formula and Distribution under H_0

$$S = \sum_{i=1}^{n} \sum_{j=1}^{R_i} \frac{1}{n - j + 1}$$

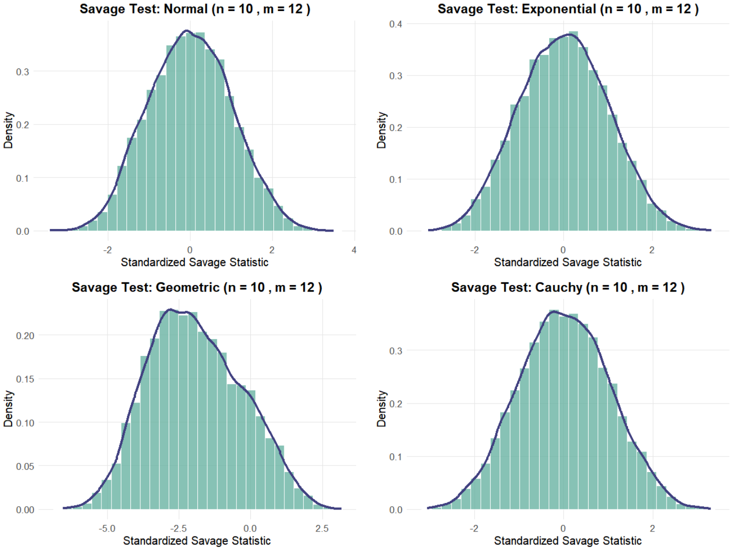
$$\mathbb{E}_{H_0}(S) = n$$

$$\mathrm{Var}_{H_0}(S) = \frac{nm}{N-1} \left(1 - \frac{1}{N} \sum_{i=1}^{N} \frac{1}{i} \right)$$

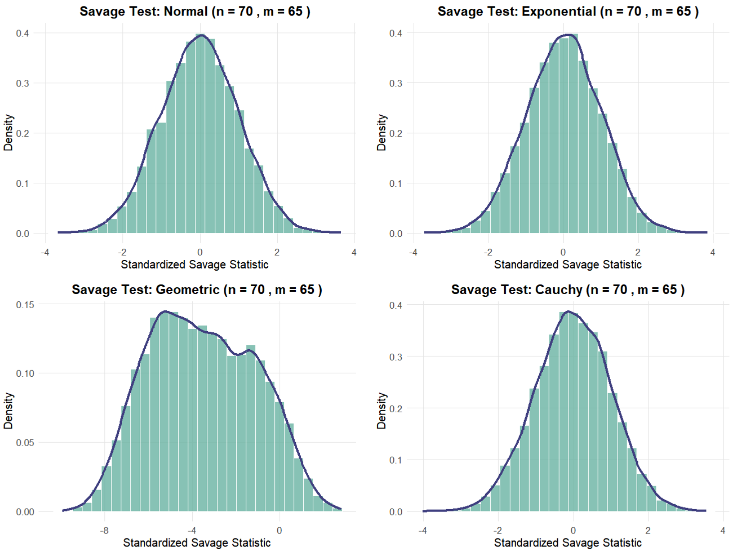
Under H_0, the standardized statistic

$$\frac{S - \mathbb{E}_{H_0}(S)}{\sqrt{\mathrm{Var}_{H_0}(S)}} \sim \mathcal{N}(0, 1)$$

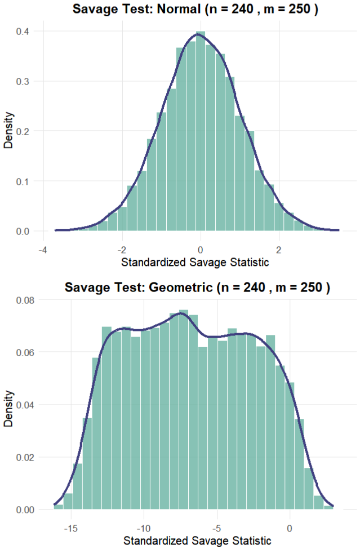
Distribution under H_0 ($n = 10$, $m = 12$)



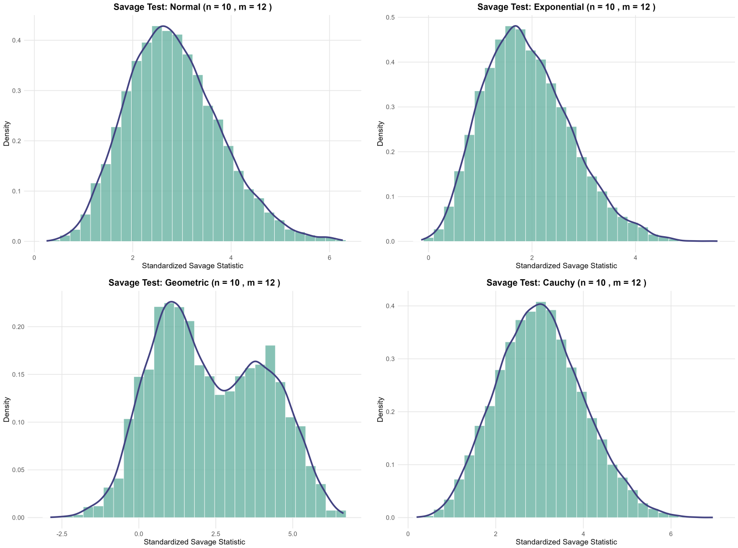
Distribution under H_0 ($n = 70$, $m = 65$)



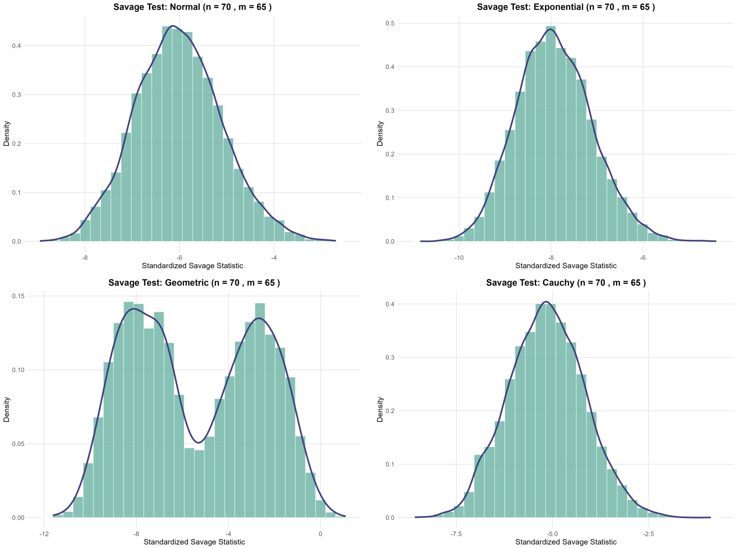
Distribution under H_0 ($n = 240$, $m = 250$)



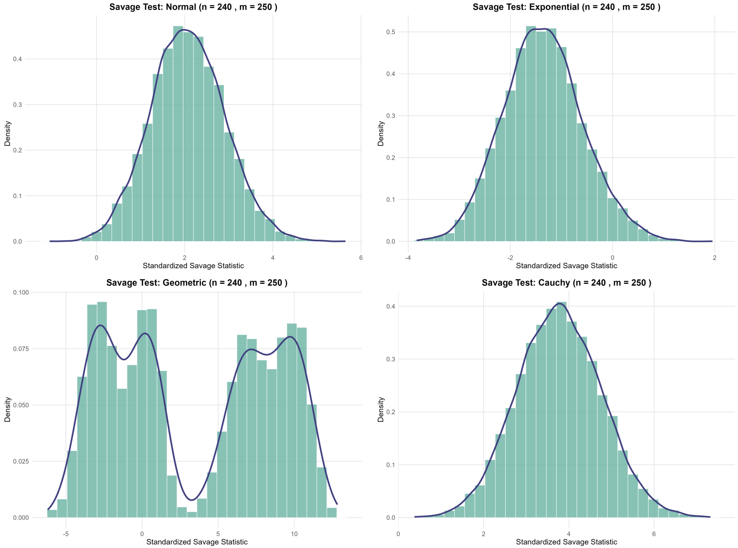
Distribution when $\theta = 2$ (n = 10, m = 12)



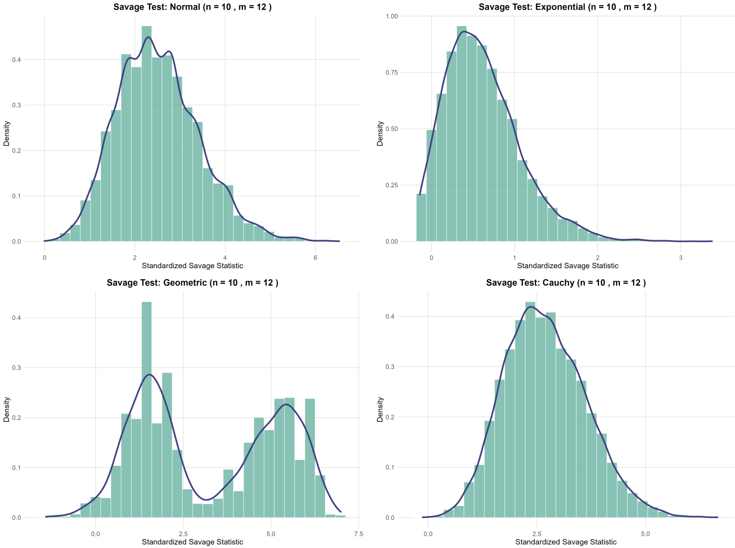
Distribution when $\theta = 2$ (n = 70, m = 65)



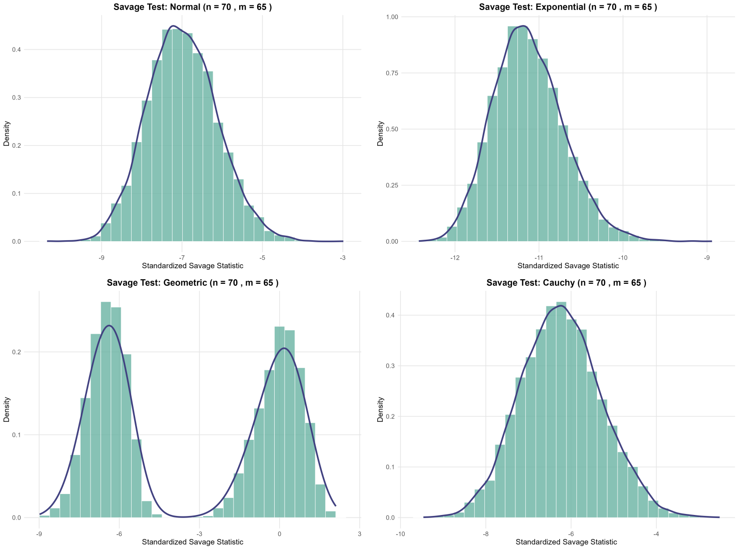
Distribution when $\theta = 2$ (n = 240, m = 250)



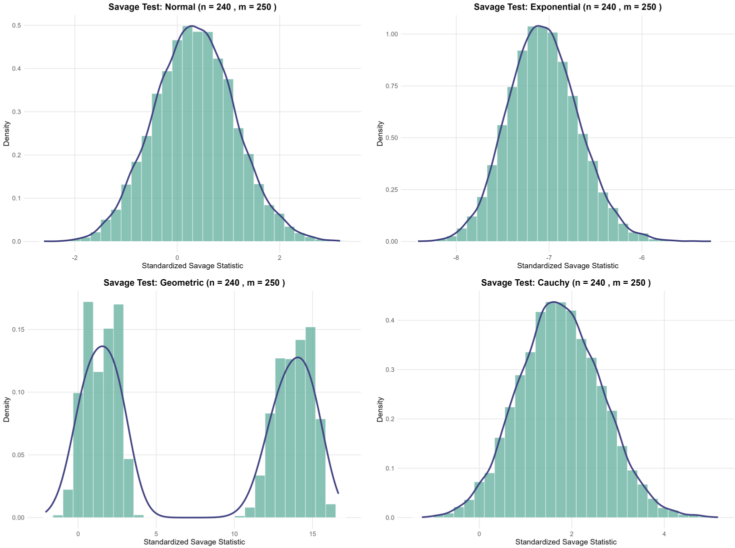
Distribution when $\theta = 6$ (n = 10, m = 12)



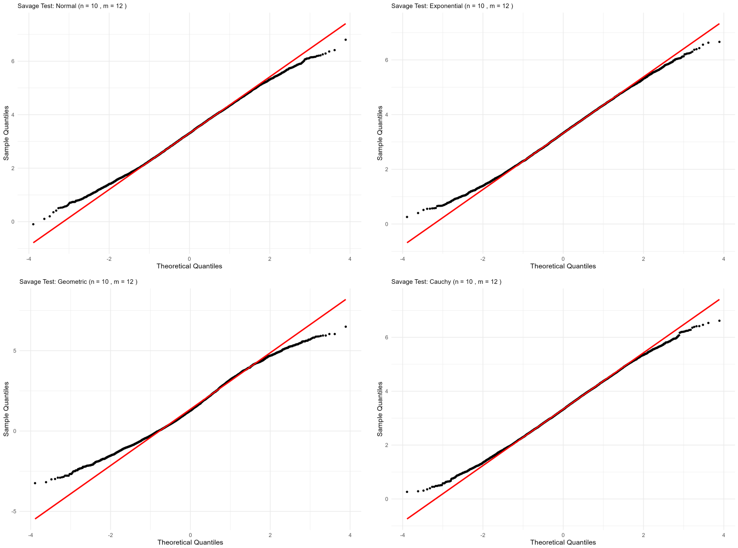
Distribution when $\theta = 6$ (n = 70, m = 65)



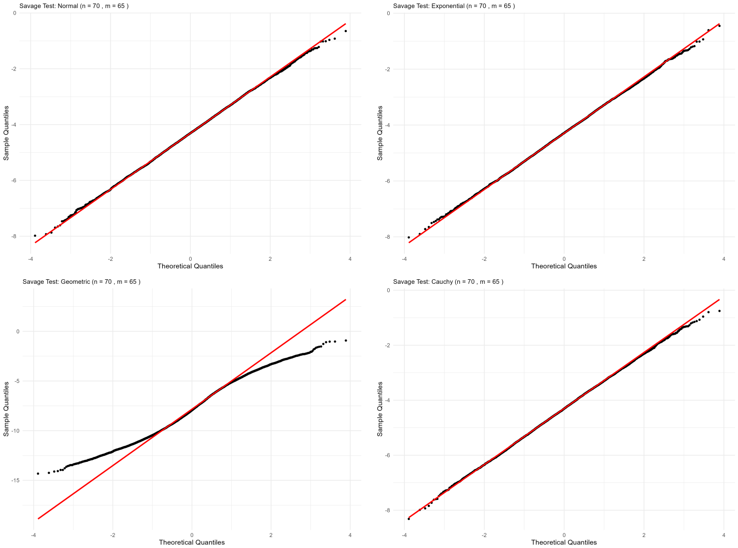
Distribution when $\theta = 6$ (n = 240, m = 250)



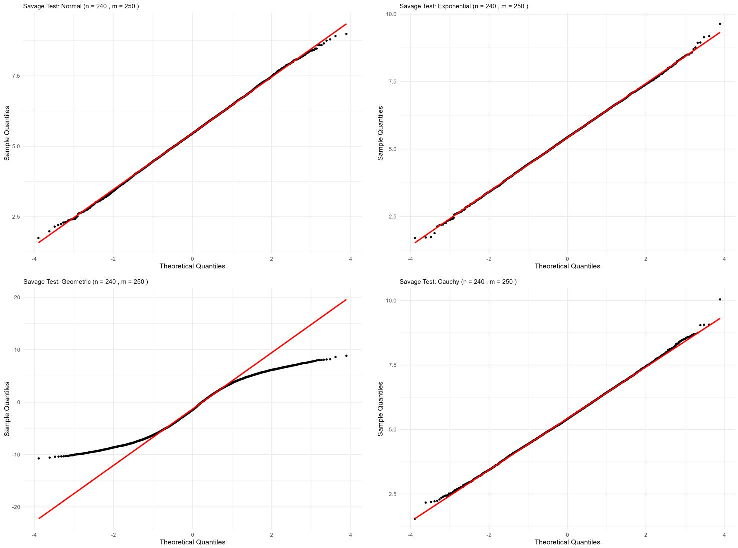
Testing for Normality under H_0 ($n = 10$, $m = 12$)



Testing for Normality H_0 (n = 70, m = 65)



Testing for Normality H_0 ($n = 240$, $m = 250$)



Asymptotic Size

Asymptotic Size Comparison for Normal Distribution ($\alpha = 0.05$)

Test	$n = m = 20$	$n = m = 50$	$n = m = 100$	$n = m = 200$
Klotz	0.0466	0.0493	0.0488	0.0508
Mood	0.0463	0.0493	0.0502	0.0537
Savage	0.0445	0.0490	0.0517	0.0500

- All tests maintain size close to nominal $\alpha = 0.05$ level
- Mood test shows slight upward trend with increasing n
- Savage test demonstrates most consistent performance

Asymptotic Size Comparison for Exponential Distribution ($\alpha = 0.05$)

Test	$n = m = 20$	$n = m = 50$	$n = m = 100$	$n = m = 200$
Klotz	0.0507	0.0503	0.0465	0.0504
Mood	0.0511	0.0500	0.0451	0.0492
Savage	0.0470	0.0456	0.0529	0.0510

- All tests show excellent size control near $\alpha = 0.05$
- Klotz and Mood maintain remarkable consistency across sample sizes
- Savage exhibits slight variability but remains within expected range

Asymptotic Size Comparison for Cauchy Distribution ($\alpha = 0.05$)

Test	$n = m = 20$	$n = m = 50$	$n = m = 100$	$n = m = 200$
Klotz	0.0479	0.0513	0.0495	0.0528
Mood	0.0504	0.0535	0.0518	0.0496
Savage	0.0480	0.0457	0.0492	0.0485

- All tests maintain proper size control under Cauchy distribution
- Klotz shows slight tendency to over-reject with larger samples
- Mood exhibits most variability across sample sizes
- Savage demonstrates most stable performance
- Results confirm robustness of nonparametric tests for heavy-tailed distributions

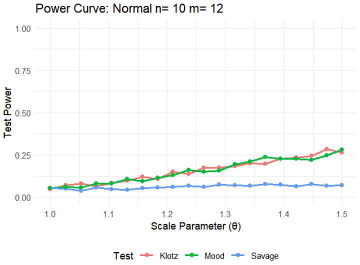
Asymptotic Size Comparison for Geometric Distribution ($\alpha = 0.05$)

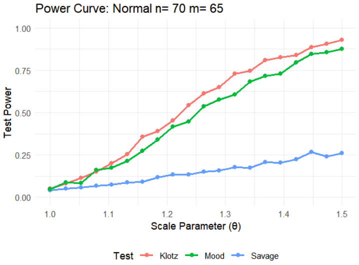
Test	$n = m = 20$	$n = m = 50$	$n = m = 100$	$n = m = 200$
Klotz	0.2550	0.6160	0.9270	0.9993
Mood	0.1454	0.3507	0.6579	0.9385
Savage	0.5745	0.6938	0.7530	0.8133

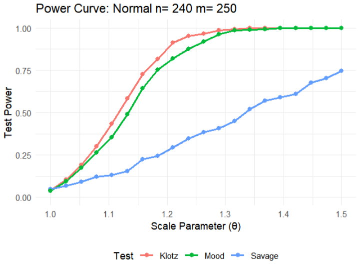
- **Severe size distortion** across all tests (nominal $\alpha = 0.05$)
- Klotz shows most extreme inflation (99.93% rejection at $n = m = 200$)
- Mood starts most conservative but escalates rapidly
- Savage maintains relatively stable (but still extreme) inflation
- **Warning**: These tests are **not valid** for Geometric data

Asymptotic Power

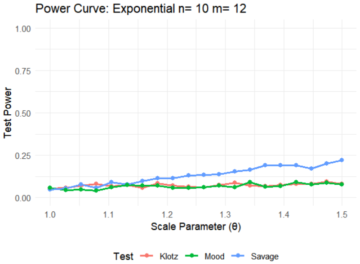
Power Curve when parent distribution is Normal

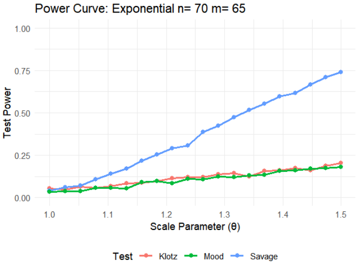


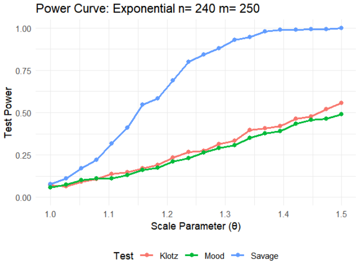




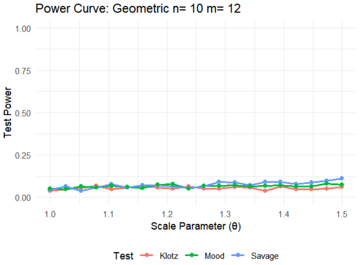
Power Curve when parent distribution is Exponential

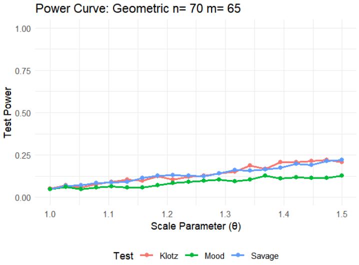


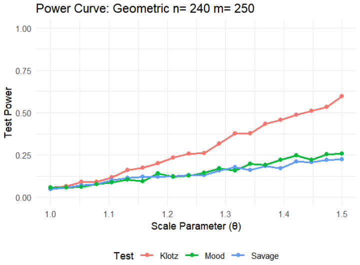




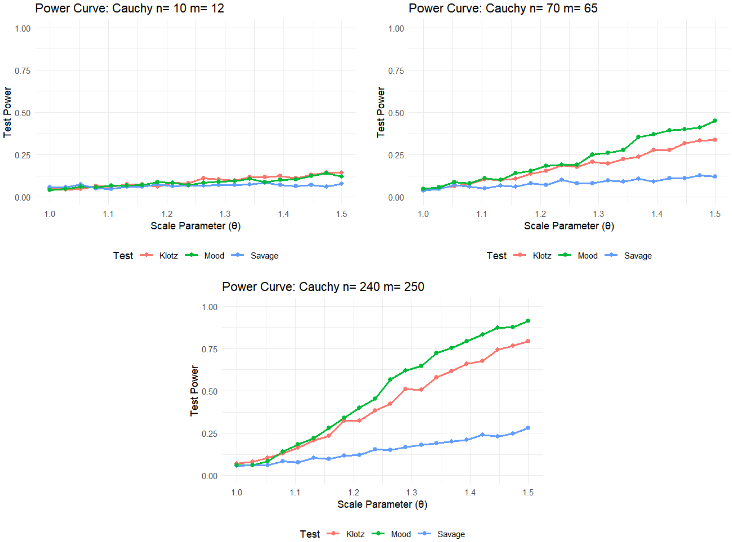
Power Curve when parent distribution is Geometric







Power Curve when parent distribution is Cauchy



Modified Discrete Rank Test

Innovative Approach for Geometric Distribution Scale Testing

Modified Discrete Rank Test (MDRT)

Traditional rank-based scale tests (e.g., Klotz, Mood, Savage) perform poorly for discrete distributions like the Geometric due to:

Problems with Classical Rank Tests

1. Assumption of Continuity:
Standard rank tests assume continuous data (no ties). However, Geometric data are discrete $\Rightarrow$ frequent ties $\Rightarrow$ inflated Type I error.

2. Improper Variance Handling:
Classical methods assume variance changes linearly with scale. But for $X \sim \text{Geometric}(p)$:

$$\text{Var}(X) = \frac{1-p}{p^2}$$

which is nonlinear in p.

Mathematical Foundations of MDRT – I

1. Variance-Stabilizing Transformation (VST)

For $X \sim \text{Geometric}(p)$, the variance is:

$$\text{Var}(X) = \frac{1-p}{p^2}$$

Since this is nonlinear in p, we apply Anscombe transformation to stabilize the variance:

$$T(X) = \sqrt{(}X + \frac{3}{8})$$

This leads to approximately constant variance:

$$\text{Var}(T(X)) \approx \text{constant} + O\left(\frac{1}{n}\right)$$

.

Mathematical Foundations of MDRT

2. Modified Rank Scores

Classical rank tests use raw ranks R_i, but MDRT uses probability scores:

$$p_i = \frac{R_i - 0.5}{N}, \quad \text{where } N = n + m$$

This transformation ensures:

$$P(p_i \leq t) \approx t, \quad \text{for } t \in (0, 1)$$

Why the -0.5 *adjustment?*
It mimics the expected values of order statistics from a continuous uniform distribution.

This adjustment improves the test's validity for discrete data.

Mathematical Foundations of MDRT

3. Test Statistic Construction

The MDRT test statistic is:

$$S = \sum_{i=1}^{n} p_i$$

Under the null hypothesis H_0 (equal scale parameters):

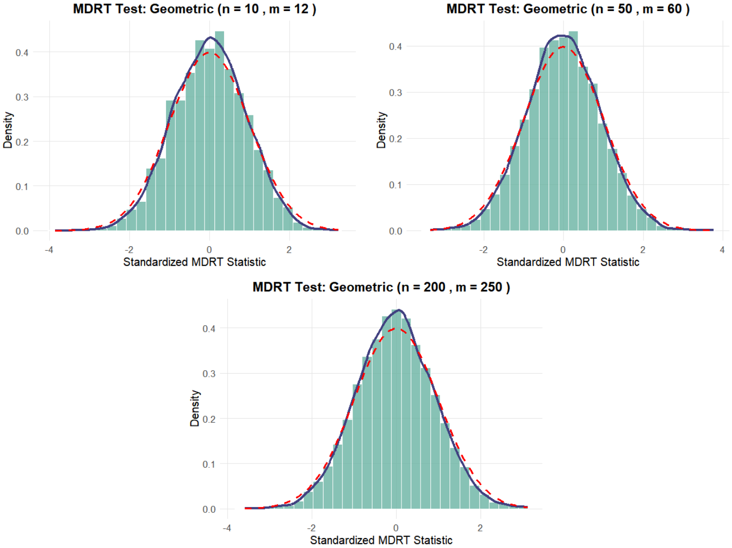
$$\mathbb{E}[S] = \frac{n}{2}, \quad \mathrm{Var}(S) = \frac{mn}{12N}\left(1 + \frac{1}{N+1}\right)$$

Standardized Form:

$$Z = \frac{S - \mathbb{E}[S]}{\sqrt{\mathrm{Var}(S)}} \xrightarrow{d} \mathcal{N}(0, 1)$$

This allows us to apply the standard normal approximation for inference in large samples.

Distribution of statistic for Geometric parent distribution



Asymptotic Size Comparison for Geometric Distribution ($\alpha = 0.05$)

$n = m = 20$	$n = m = 50$	$n = m = 100$	$n = m = 200$
0.0324	0.0348	0.0348	0.0333

- The test maintains appropriate control near nominal level across all n

Comparison of Power Curves