



UNIVERSITÀ DEGLI STUDI DI MILANO

DIPARTIMENTO DI SCIENZE CLINICHE
E DI COMUNITÀ

Scuola di Dottorato in Scienze Biomediche Cliniche e Sperimentali
Corso di Dottorato di Ricerca in Statistica Biomedica
Settore Scientifico-Disciplinare MED/01 - XXVI CICLO

Tesi di Dottorato di Ricerca

IDENTIFICAZIONE DEGLI ASSI FATTORIALI INFORMATIVI
NELL'ANALISI DELLE COMPONENTI PRINCIPALI

DOTTORANDA: Maddalena Plebani (R09128)

RELATORE: Prof. Silvano Milani

COORDINATORE DEL CORSO DI DOTTORATO: Prof. Adriano Decarli

ANNO ACCADEMICO 2012/2013

Sommario

1.	Introduzione	2
2.	Regole empiriche per la PCA con matrici di correlazione: risultati dalla letteratura e razionale tesi...	4
3.	Metodi e risultati	7
3.1	La generazione dei dati	7
3.2	Calcolo dei valori soglia in funzione della dimensione dello studio e del numero di variabili: metodo diretto	8
3.2.1	Funzione per il calcolo del 95° centile	8
3.2.2	Rischio d'errore di tipo I e probabilità condizionate	10
3.2.3	Adeguatezza della funzione	14
3.3	Calcolo dei valori soglia in funzione della dimensione dello studio e del numero di variabili: metodo indiretto	15
3.3.1	Funzione per il calcolo del 95° centile	15
3.3.2	Interpolazione dei parametri μ , ξ , $\bar{\lambda}$ in funzione di n , k	16
3.3.3	Rischio d'errore di tipo I e probabilità condizionate	18
3.3.4	Adeguatezza della funzione	24
4.	Discussione e conclusioni	26
5.	Appendice	28
6.	Bibliografia	80

1. Introduzione

L'analisi delle componenti principali (*principal component analysis*, PCA) è una tecnica statistica introdotta all'inizio del secolo scorso da Karl Pearson e successivamente sviluppata da Hotelling (1933) con lo scopo di trasformare un set di k variabili di partenza in un altro set di $q < k$ variabili; queste nuove variabili sono combinazioni lineari delle variabili originarie che massimizzano a scalare la varianza spiegata, sotto il vincolo di non correlazione.

Questa tecnica permette di rappresentare fenomeni anche complessi attraverso l'utilizzo di un minor numero di variabili, il cui significato è stabilito *a posteriori* attraverso l'analisi delle correlazioni tra le componenti principali e le variabili originarie; in alternativa, le componenti principali possono essere pensate come *proxy* di fattori latenti (approccio *a priori*; in questo caso, la PCA si riconduce alla più generale analisi dei fattori. Ricci e Milani 2007).

Importanza cruciale nella PCA è rivestita proprio dal criterio per la determinazione del numero di fattori da considerare come espressione di un costrutto latente sottostante. A questo scopo è utile valutare le distribuzioni degli autovalori della matrice di covarianza o correlazione in esame. Questo compito può essere condotto utilizzando un approccio matematico o un approccio empirico.

Tuttavia, affrontare il problema da un punto di vista analitico-matematico trovando la distribuzione degli autovalori di una matrice di covarianza o correlazione può avere risvolti oltremodo complessi. Molti studi di questo tipo si concentrano soltanto sulle matrici di covarianza (meglio trattabili dal punto di vista matematico rispetto alle matrici di correlazione) e presentano risultati di natura asintotica derivati da particolari assunti. Le distribuzioni degli autovalori possono essere infatti sviluppate sia in termini esatti che in termini asintotici; tuttavia, le formule per esprimere le distribuzioni esatte sono spesso analiticamente poco trattabili, o convergono molto lentamente, mentre il rispettivo comportamento al limite può essere espresso attraverso funzioni più semplici o calcolabili con minori difficoltà (si veda Muirhead, 1978, pp. 5-6; Muirhead, 1982, p. 390). Inoltre, essendo il caso complesso matematicamente più facile da trattare, parecchi risultati ottenuti sono limitati a quest'ambito.

Per quanto riguarda la matrice di correlazione, il problema della distribuzione limite degli autovalori di una matrice di correlazione è stato inizialmente trattato da Anderson (1963); un compendio riassuntivo dello stato dell'arte della ricerca alla fine degli anni Settanta lo si può trovare in Konishi (1979); una revisione più aggiornata degli studi sulle matrici di correlazione si ha successivamente in Boik (2003, 2013), che propone modelli flessibili per le componenti principali (2003, 2004, 2013). Novità sono contenute anche in Jiang (2004), il quale mostra la convergenza quasi certa dell'autovalore più grande e di quello più piccolo di una matrice di correlazione. Sotto particolari assunti, la distribuzione asintotica dell'autovalore maggiore o minore opportunamente riscalati di una matrice di correlazione è stata poi successivamente

ottenuta da Bao *et al.* (2011) e consiste nella distribuzione di Tracy-Widom. Infine, nuovi risultati asintotici universali con matrici di correlazione sono contenuti in Pillai e Yin (2012).

Per ovviare alle difficoltà riscontrate a livello analitico-matematico, in letteratura molti studi hanno sviluppato e utilizzato diversi metodi empirici. Alcuni di questi metodi sono basati su argomenti teorici, quali la regola di Kaiser-Guttman, che prevede di considerare solo quelle componenti il cui autovalore sia maggiore di 1. Altri sono metodi grafici, quali lo *scree plot* o il metodo *broken-stick*. Tuttavia queste regole nel loro complesso sono spesso caratterizzate da basi matematiche poco solide, oppure non sono sempre correttamente applicate. I principali risultati forniti dalla letteratura riguardanti regole empiriche per l'individuazione di assi fattoriali informativi nella PCA a partire da matrici di correlazione sono presentati nel paragrafo 2.

L'interesse alla risoluzione di questo problema è trasversale a più ambiti di ricerca: oltre a essere un'intrigante questione di teoria matematica, è stato e ed è tuttora affrontato sia a livello teorico che a livello empirico dalla fisica quantistica, dalla meccanica, dal *signal processing*, dalla finanza, dalla psicologia. In campo statistico, tuttavia, l'interesse principale è applicativo e risiede nell'uso della PCA, solitamente applicata a matrici di correlazione anziché di covarianza, per trattare realtà complesse descritte da più variabili, distribuite secondo funzioni differenti, sulla base di dataset anche di piccole dimensioni, per i quali non valgono le regole asintotiche. Come esplicitato da Johnstone (2006, p. 322): "[...] we very much need extensions of the distributional results to real data [...]". A questo scopo, nella presente tesi si sono generati per simulazione 6001 campioni per ognuno dei 70 scenari previsti, caratterizzati da diverse dimensioni dello studio (n) e un diverso numero di variabili (k), tutte distribuite come normali standard (par. 3.1), sotto l'assunto che non esistano fattori latenti. Gli autovalori sono stati calcolati a partire dalle matrici di correlazione. L'argomento in esame è stato trattato sotto due differenti metodi: diretto (par. 3.2) e indiretto (par. 3.3). Nel primo caso, il valore del 95° centile della distribuzione degli autovalori è stato espresso in funzione di n e k . Nel secondo caso, la *distribuzione* di ciascuno dei primi quattro autovalori è stata ricondotta a una Normale standard attraverso l'applicazione della trasformazione Box e Cox, e i parametri della trasformazione sono stati espressi in funzione di n e k . Nei relativi sottoparagrafi si è valutata l'adeguatezza delle rispettive funzioni. In entrambi i casi sono stati considerati soltanto i primi 4 autovalori. Nell'Appendice sono contenuti i grafici relativi alla distribuzione degli autovalori, normalizzata e non, per i vari scenari (come da par. 3.3).

2. Regole empiriche per la PCA con matrici di correlazione: risultati dalla letteratura e razionale tesi

In letteratura sono proposte varie regole empiriche per l'individuazione degli assi fattoriali informativi nell'Analisi delle Componenti Principali.

Uno di questi metodi è denominato Parallel Analysis (PA). La PA ha i suoi albori nel 1965, con un lavoro di Horn pubblicato su *Psychometrika* che mostra empiricamente quanto possano tra loro differire gli autovalori campionari e della popolazione e come regole quali quella di Kaiser (considerare i soli autovalori maggiori di 1) possano indurre a scelte sbagliate sul numero di componenti da considerare come informative. Il metodo prevede un approccio empirico al problema del numero di componenti/fattori considerati rilevanti nell'analisi delle componenti principali/dei fattori, e consiste nel simulare attraverso il metodo Monte Carlo delle realizzazioni di certe variabili casuali dati i parametri n e k , e confrontare i risultati ottenuti con quelli osservati a partire dalla matrice di correlazione. I valori soglia proposti da Horn sono quelli attesi; successivamente, Longman *et al.* (1989) e Glorfeld (1995) hanno proposto valori soglia corrispondenti al 95° centile; questi ultimi hanno poi preso il sopravvento in letteratura.

Timmerman e Lorenzo-Seva (2011) descrivono alcuni programmi utilizzati per eseguire la PA. FACTOR (Lorenzo-Seva e Ferrando, 2006) è un programma *standalone* utile per l'approccio PA nella PCA, così come il software VISTA-Paran (Valerio-Mora 2007). Anche il programma descritto in Watkins (2010) serve allo stesso scopo, ma utilizza le medie come riferimento, anziché i 95° centili. L'approccio PA non è tuttavia disponibile "di serie" per i più diffusi software commerciali. O'Connor (2000) ha pubblicato codici per eseguire la PA in ambiente SPSS, SAS e MATLAB, codici scaricabili direttamente dal link: <https://people.ok.ubc.ca/briocconn/nfactors/nfactors.html>). Un'altra macro per SAS è disponibile ai link: <http://www2.sas.com/proceedings/sugi28/090-28.pdf> e <http://www.mrg.com/articles/parallel.sas> (Kabacoff 2000). Altri codici per SPSS sono contenuti in Hayton *et al.* (2004) e in Thompson e Daniel (1996). Il codice in R è invece disponibile all'indirizzo: <http://cran.r-project.org/web/packages/paran/paran.pdf>; Dinno 2013). Il programma *paran* scritto in STATA pubblicato da Dinno (2009b) è disponibile all'indirizzo <http://www.doyenne.com/stata/>. Un'interfaccia online per il calcolo del 95° centile è disponibile all'indirizzo <http://ires.ku.edu/~smishra/parallelengine.htm> (Patil 2008).

Approcci alternativi alla Parallel Analysis, che ricorre al metodo Monte Carlo per la generazione di dati, sono basati sul metodo permutativo (per esempio, Peres-Neto *et al.* 2004) e sul metodo bootstrap (Gebert e Ferreira 2013; Larsen e Warne, 2010, Babamoradi *et al.*, 2013; per il metodo jackknife sulle matrici di correlazione: Nagao, 1987). Un ulteriore approccio consiste nell'utilizzo della cross-validation

(Wold, 1976 e 1978; Eastment e Krzanowski 1982; Josse e Husson 2012, anche per una revisione del metodo).

Per quanto riguarda l'approccio permutativo, Buja e Eyuboglu (1992) hanno mostrato attraverso simulazioni che il metodo della parallel analysis risulta essere robusto, e che perciò la permutazione è di limitata importanza pratica (p. 509): "We give evidence [...] that quasi-inferential PA based on normal random variates (as opposed to data permutations) is surprisingly independent of distributional assumptions [...]. This is a justification for providing tables for quasi-inferential PA". Analogo risultato è stato ottenuto da Hayton *et al.* (2004, 2009) e Dinno (2009a): la Monte Carlo Parallel Analysis risulta essere una tecnica robusta, non sensibile a cambiamenti della distribuzione sottostante (si veda anche Zorić e Opačić 2013). Analoghi risultati sono stati ottenuti da Peres-Neto *et al.* (2004), i quali hanno mostrato quanto il metodo bootstrap non sia preferibile al metodo di generazione dei dati da una particolare distribuzione, nel caso dello scenario a 0 fattori latenti.

Critiche all'approccio della Parallel Analysis riguardano invece la valutazione degli autovalori successivi al primo effettuate con questo metodo; evidenziano infatti Buja e Eyuboglu (1992) e Turner (1998), nelle parole di Hayton *et al.* (2004, p. 202): "One caution about PA is offered by Turner (1998), who noted that due to the interdependent nature of eigenvalues, the presence of a large first factor in a PA will reduce the size of [the remaining] noise eigenvalues. The consequence is that in certain situations, PA can underfactor, which is potentially more serious than overfactoring. The impact of this limitation is most serious for smaller sample sizes, where there is a high correlation between factors, or where a second factor is based on a relatively small number of items. However, Turner also noted that when used with care, PA remains a potentially "valuable method [...]" ". Crawford *et al.* (2010, p. 888) chiariscono, sempre riferendosi al metodo della Parallel Analysis: "Theoretically, this rationale may be legitimate for evaluation of the first eigenvalue, as the empirical distribution of the first eigenvalue from the simulated data represents the sampling distribution under the null hypothesis that there are no common factors. However, this rationale is questionable for evaluating the remaining eigenvalues, as the empirical distribution of the k th eigenvalue ($k > 1$) represents the hypothesis that there are no common factors rather than the desired hypothesis that there are $k - 1$ factors. Accordingly, PA must be viewed as a heuristic approach rather than a mathematically rigorous approach resting on a solid conceptual base.". Ribadisce Turner (1998, p. 544): "PA sets up a null hypothesis to test if the first eigenvalue is caused by a true factor and that null hypothesis is appropriate only for the first eigenvalue. One of the main problems with PA is that the tables, formulas or programs that have been designed to facilitate PA all assume that real data eigenvalues can be compared directly to eigenvalues determined from data in which no real factors are present".

Per una rivisitazione della Parallel Analysis per autovalori successivi al primo si veda Green *et al.* (2012) e in particolare gli Studi 1 e 2 condotti e descritti in Turner 1998 (p. 550 e p. 556), i quali lo inducono a concludere come segue (p. 565): "To make a PA more accurate, the researcher should not only mimic the sample size and the number of items but should also model into the analysis the size and the structure pattern of known real factors. This would require a series of simulations, each testing the hypothesis that the next eigenvalue is due to chance".

Oltre ad articoli che descrivono regole empiriche per la determinazione del numero di fattori, negli ultimi anni si sono susseguiti vari studi il cui scopo era quello di procedere ad una comparazione delle varie regole empiriche; tra questi emergono i lavori di Zwick e Velicer (1986); Jackson *et al.* (1993), Velicer *et al.* (2000), Peres-Neto *et al.* (2004). Molti di questi studi hanno constatato come sia più efficiente adottare più di una regola per verificare quale sia effettivamente il numero di fattori latenti da considerare (si vedano in proposito: Silverstein 1990, Eaton *et al.* 1999; Zwick and Velicer 1986, Velicer *et al.* 2000, Turner 1998, Hayton *et al.* 2004, Fabrigar *et al.* 1999, Ford *et al.* 1986, Coste *et al.* 2005), anche poiché spesso la regola più efficiente per uno scenario è soppiantata da un'altra regola quando lo scenario muta. Tuttavia, spesso in questi lavori il metodo della Parallel Analysis figura come uno tra i più accurati. Se ci si limita al caso analizzato nella presente tesi, ossia lo scenario a 0 fattori latenti, molte ricerche hanno consigliato l'utilizzo della Parallel Analysis (tra gli altri: Buja e Eyuboglu 1992; Glorfeld 1995, Crawford 2010).

Un altro metodo presente in letteratura per l'individuazione di assi fattoriali informativi nella PCA è rappresentato da equazioni di regressione. Il vantaggio di un approccio simile si trova ben esplicitato in Velicer *et al.* (2000, p. 50): "[...] much recent work has been focused on developing alternatives to avoid the necessity of generating multiple correlation matrices. One alternative is the development of regression equations for predicting the random eigenvalues."

Come evidenziato da Lautenschlager (1989b) e da Silverstein (1990), nel passato equazioni di regressione per la predizione di valori soglia per gli autovalori nella PCA sono state utilizzate da Allen e Hubbard (1986), Lautenschlager *et al.* (1989a), Longman *et al.* (1989), Keeling (2000). I primi due lavori presentano formule ricorsive e modellano le medie, mentre a partire da Longman *et al.* (1989) le formule adottate non sono più ricorsive e inoltre è presa in considerazione non solo la media come soglia, ma anche il 95° centile, e per entrambe le soglie sono proposte apposite equazioni; il lavoro di Keeling (2000) è ammirevole per la proposta di equazioni non corredate da tabelle per i vari coefficienti a seconda dell'autovalore considerato; tuttavia non presenta risultati relativi al 95° centile, bensì alle medie.

Critiche al metodo di regressione provengono da Dinno (2009a), la quale cita a sostegno della sua tesi i lavori di Velicer (2000) e Lautenschlager (1989b). Tuttavia il lavoro di Lautenschlager prende in considerazione le medie, non i 95° centili per differenti scenari; inoltre Velicer riadisce (p. 13) “For the three regression equation approaches and the tabled value approach, there has not been a study that evaluates the ability of the method to accurately determine m .” (m rappresenta il numero di fattori da considerare come espressione di un costrutto latente).

Conferme al metodo di regressione provengono invece da Cota *et al.* (1993b; si veda in particolare la Tabella 1 a p. 871): è probabile che si pervenga allo stesso risultato adottando il metodo della regressione secondo Longman *et al.* (1989a) o il metodo dell'interpolazione lineare dei valori del 95° centile riportati in tabelle divisi per scenari.

Lo scopo di questa tesi è illustrare, nell'ambito della PA, due tecniche per la derivazione di valori soglia della distribuzione dei primi autovalori della matrice di correlazione in funzione soltanto della numerosità campionaria (n) e del numero di variabili (k), così da evitare il ricorso a tabelle estensive di valori soglia per i diversi scenari o la generazione di dati con procedure di calcolo intensivo (quali i metodi Monte Carlo, permutazionali, bootstrap, la cross-validation o altro ancora).

3. Metodi e risultati

3.1 La generazione dei dati

La generazione dei dati è stata condotta considerando 70 diversi scenari, ottenuti combinando 7 diversi valori per n (75, 150, 300, 600, 1200, 2400, 4800) con 10 diversi valori per k (6, 12, 18, 24, 30, 36, 48, 54, 60). I dati sono stati generati con l'ausilio del software SAS® (v.9.2, 2008), con le funzioni STREAMINIT e RAND(“NORMAL”)¹, sotto gli assunti che le variabili si distribuiscano come Normali standard univariate e che non esistano fattori latenti (ossia che le variabili siano indipendenti). Le matrici di correlazione su cui effettuare la PCA sono state dunque calcolate su campioni tratti da una distribuzione gaussiana multivariata con matrice di correlazione (P) coincidente con la matrice di identità (I).

L'indipendenza è garantita dall'algoritmo di Matsumoto-Nishimura, su cui si basa la funzione RAND, dotato di un periodo dell'ordine di 10^{6001} (Matsumoto e Nishimura 1998). Si sono così generati

¹ Risultati quasi identici sono ottenibili attraverso l'utilizzo di SAS/IML (v.9.2, 2008), utilizzando rispettivamente le funzioni RANDSEED e RANDNORMAL.

3,429,571,500² valori con distribuzione Normale standard. In quel che segue sono stati presi in considerazione esclusivamente i primi 4 autovalori.

3.2 Calcolo dei valori soglia in funzione della dimensione dello studio e del numero di variabili: metodo diretto

3.2.1 Funzione per il calcolo del 95° centile

La PCA è stata applicata ad ognuno dei 6001 campioni di ognuno dei 70 scenari a partire dalla matrice di correlazione. Per ogni scenario si è presa in considerazione l'osservazione corrispondente al 95° centile della distribuzione di ciascuno dei primi 4 autovalori (5701^a osservazione). Il 95° centile della distribuzione dei primi 4 autovalori può essere predetto dalla seguente funzione non lineare a 7 parametri in n e k :

$$l_{95^\circ} = 1 + (\beta_1 n^{\beta_2}) k^{(\gamma_0 + \gamma_1 n^{\gamma_2})} + \delta (nk)^\eta$$

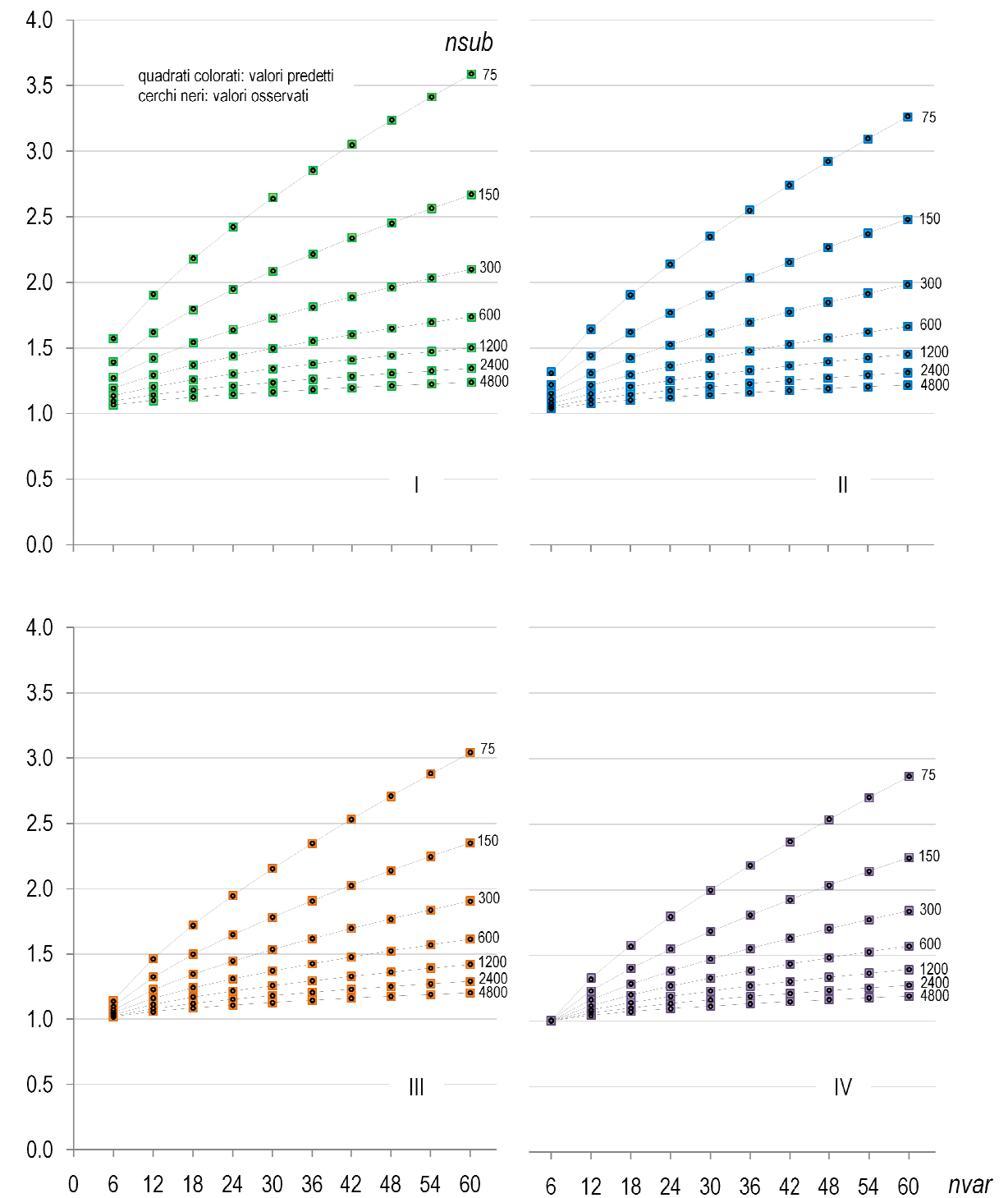
I valori dei 7 parametri sono stati ottenuti con il metodo dei minimi quadrati pesati, usando come peso \sqrt{nk} . Per ognuno dei 4 autovalori considerati, i valori dei parametri così ottenuti sono riportati in Tabella 1. I grafici qui di seguito mostrano come i valori del 95° centile predetti dalla funzione approssimino quelli osservati (Figure 1-4).

Tabella 1. Valore dei parametri (metodo diretto)

autovalore	valori numerici dei parametri						
	β_1	β_2	γ_0	γ_1	γ_2	δ	η
I	1.1303	-0.4204	0.4163	0.6476	-0.2394	-0.0192	-0.0389
II	0.6799	-0.3777	0.3019	0.7481	-0.1470	-1.4500	-0.3740
III	0.4904	-0.3596	0.2230	0.8356	-0.1119	-3.2728	-0.4230
IV	0.3598	-0.3351	0.0615	1.0183	-0.0838	-4.6723	-0.4321

² $(75+150+300+600+1200+2400+4800) \times 6001 \times 60 = 3,429,571,500$.

Figura 1-4. 95° centili predetti con il metodo diretto e 95° centili osservati per ogni scenario (I-IV autovalore)



3.2.2 Rischio d'errore di tipo I e probabilità condizionate

La Tabella 2 riporta i rischi di errore di tipo I (α) calcolati sul complesso dei 70 scenari considerando come soglie i valori osservati del 95° centile della distribuzione degli autovalori e quelli predetti dalla funzione. Con tali valori soglia, il rischio α è, ovviamente, del 5% per ciascuno degli autovalori, mentre la probabilità che almeno i primi p ($p = 2, 3, 4$) autovalori superino tutti la rispettiva soglia è maggiore di 0.05^p, come ci si attenderebbe sotto assunto di non correlazione tra gli autovalori. In realtà, le coppie di autovalori consecutivi sono sempre *positivamente correlate* (Figure 5-11), e il valore della correlazione cresce all'aumentare di k e, anche se in misura minore, di n .

Le Tabelle 3 e 4 riportano, per ciascuno dei 70 scenari e per ognuno dei 4 autovalori, le probabilità *congiunte* ottenute utilizzando le soglie osservate e quelle predette con il metodo diretto: le probabilità più basse corrispondono agli scenari in cui le correlazioni tra primo e secondo autovalore sono inferiori (in particolare gli scenari a 6 variabili), mentre le probabilità più alte sono quelle corrispondenti a scenari con un numero di variabili superiore a 30. Nella Tabella 5 sono invece riportate le differenze tra la probabilità *condizionata* di superamento della soglia osservata e la probabilità condizionata di superamento della soglia predetta con il metodo diretto; queste tabelle contengono solo gli scenari dove tali probabilità mostrano la massima differenza (positiva o negativa) per almeno un autovalore.

Tabella 2. Frequenza assoluta e percentuale di superamento delle soglie osservate e predette con il metodo diretto (media su tutti gli scenari)

	soglie osservate		soglie predette	
	freq	%	freq	%
<i>almeno il 1° autovalore > 95°</i>	21000	5.00	21021	5.00
<i>almeno i primi 2 > 95°</i>	3418	0.81	3451	0.82
<i>almeno i primi 3 > 95°</i>	800	0.19	794	0.19
<i>almeno i primi 4 > 95°</i>	217	0.05	218	0.05

Tabella 3. Probabilità congiunte calcolate con le soglie osservate per ogni scenario

scenario		probabilità congiunte				scenario		probabilità congiunte			
<i>n</i>	<i>k</i>	<i>almeno il 1° autovalore > 95°</i>	<i>almeno i primi 2 > 95°</i>	<i>almeno i primi 3 > 95°</i>	<i>almeno i primi 4 > 95°</i>	<i>n</i>	<i>k</i>	<i>almeno il 1° autovalore > 95°</i>	<i>almeno i primi 2 > 95°</i>	<i>almeno i primi 3 > 95°</i>	<i>almeno i primi 4 > 95°</i>
75	6	0.0500	0.0032	0.0000	0.0000	600	36	0.0500	0.0093	0.0018	0.0007
75	12	0.0500	0.0042	0.0005	0.0000	600	42	0.0500	0.0113	0.0022	0.0005
75	18	0.0500	0.0050	0.0008	0.0000	600	48	0.0500	0.0088	0.0023	0.0010
75	24	0.0500	0.0073	0.0015	0.0002	600	54	0.0500	0.0127	0.0035	0.0022
75	30	0.0500	0.0073	0.0015	0.0007	600	60	0.0500	0.0098	0.0027	0.0010
75	36	0.0500	0.0083	0.0018	0.0007	1200	6	0.0500	0.0045	0.0003	0.0000
75	42	0.0500	0.0083	0.0020	0.0005	1200	12	0.0500	0.0080	0.0005	0.0002
75	48	0.0500	0.0067	0.0012	0.0002	1200	18	0.0500	0.0068	0.0015	0.0003
75	54	0.0500	0.0088	0.0018	0.0000	1200	24	0.0500	0.0088	0.0030	0.0005
75	60	0.0500	0.0098	0.0028	0.0002	1200	30	0.0500	0.0108	0.0020	0.0007
150	6	0.0500	0.0043	0.0000	0.0000	1200	36	0.0500	0.0088	0.0022	0.0003
150	12	0.0500	0.0057	0.0012	0.0000	1200	42	0.0500	0.0072	0.0020	0.0005
150	18	0.0500	0.0073	0.0008	0.0002	1200	48	0.0500	0.0083	0.0023	0.0008
150	24	0.0500	0.0077	0.0015	0.0002	1200	54	0.0500	0.0100	0.0033	0.0005
150	30	0.0500	0.0080	0.0012	0.0003	1200	60	0.0500	0.0107	0.0038	0.0008
150	36	0.0500	0.0085	0.0025	0.0007	2400	6	0.0500	0.0043	0.0007	0.0000
150	42	0.0500	0.0077	0.0012	0.0002	2400	12	0.0500	0.0067	0.0018	0.0000
150	48	0.0500	0.0083	0.0017	0.0005	2400	18	0.0500	0.0090	0.0020	0.0003
150	54	0.0500	0.0080	0.0025	0.0007	2400	24	0.0500	0.0082	0.0025	0.0007
150	60	0.0500	0.0082	0.0020	0.0010	2400	30	0.0500	0.0088	0.0025	0.0005
300	6	0.0500	0.0057	0.0005	0.0000	2400	36	0.0500	0.0087	0.0022	0.0008
300	12	0.0500	0.0068	0.0007	0.0002	2400	42	0.0500	0.0098	0.0022	0.0007
300	18	0.0500	0.0080	0.0022	0.0007	2400	48	0.0500	0.0095	0.0033	0.0017
300	24	0.0500	0.0088	0.0017	0.0005	2400	54	0.0500	0.0098	0.0027	0.0012
300	30	0.0500	0.0093	0.0020	0.0003	2400	60	0.0500	0.0102	0.0037	0.0020
300	36	0.0500	0.0093	0.0018	0.0003	4800	6	0.0500	0.0042	0.0000	0.0000
300	42	0.0500	0.0088	0.0022	0.0007	4800	12	0.0500	0.0060	0.0012	0.0000
300	48	0.0500	0.0102	0.0035	0.0010	4800	18	0.0500	0.0073	0.0012	0.0000
300	54	0.0500	0.0078	0.0020	0.0003	4800	24	0.0500	0.0098	0.0022	0.0003
300	60	0.0500	0.0092	0.0023	0.0002	4800	30	0.0500	0.0075	0.0020	0.0008
600	6	0.0500	0.0037	0.0003	0.0000	4800	36	0.0500	0.0087	0.0020	0.0010
600	12	0.0500	0.0077	0.0013	0.0002	4800	42	0.0500	0.0103	0.0030	0.0008
600	18	0.0500	0.0082	0.0025	0.0008	4800	48	0.0500	0.0087	0.0032	0.0010
600	24	0.0500	0.0098	0.0022	0.0010	4800	54	0.0500	0.0100	0.0027	0.0008
600	30	0.0500	0.0102	0.0022	0.0005	4800	60	0.0500	0.0100	0.0032	0.0008

Tabella 4. Probabilità congiunte calcolate con le soglie predette per ogni scenario (metodo diretto)

scenario		probabilità congiunte				scenario		probabilità congiunte			
<i>n</i>	<i>k</i>	<i>almeno il 1° autovalore > 95°</i>	<i>almeno i primi 2 > 95°</i>	<i>almeno i primi 3 > 95°</i>	<i>almeno i primi 4 > 95°</i>	<i>n</i>	<i>k</i>	<i>almeno il 1° autovalore > 95°</i>	<i>almeno i primi 2 > 95°</i>	<i>almeno i primi 3 > 95°</i>	<i>almeno i primi 4 > 95°</i>
75	6	0.0498	0.0025	0.0000	0.0000	600	36	0.0490	0.0098	0.0018	0.0007
75	12	0.0548	0.0050	0.0008	0.0000	600	42	0.0495	0.0118	0.0022	0.0007
75	18	0.0528	0.0052	0.0008	0.0000	600	48	0.0503	0.0097	0.0025	0.0010
75	24	0.0497	0.0073	0.0013	0.0002	600	54	0.0542	0.0148	0.0038	0.0023
75	30	0.0452	0.0063	0.0010	0.0005	600	60	0.0532	0.0105	0.0028	0.0010
75	36	0.0498	0.0083	0.0018	0.0007	1200	6	0.0428	0.0042	0.0003	0.0000
75	42	0.0478	0.0082	0.0020	0.0005	1200	12	0.0450	0.0068	0.0003	0.0002
75	48	0.0507	0.0072	0.0012	0.0002	1200	18	0.0515	0.0080	0.0017	0.0005
75	54	0.0497	0.0088	0.0018	0.0000	1200	24	0.0502	0.0090	0.0030	0.0005
75	60	0.0518	0.0110	0.0032	0.0002	1200	30	0.0537	0.0123	0.0022	0.0007
150	6	0.0437	0.0037	0.0000	0.0000	1200	36	0.0543	0.0098	0.0022	0.0007
150	12	0.0568	0.0065	0.0013	0.0000	1200	42	0.0497	0.0077	0.0022	0.0007
150	18	0.0572	0.0100	0.0012	0.0002	1200	48	0.0497	0.0088	0.0025	0.0008
150	24	0.0533	0.0087	0.0018	0.0002	1200	54	0.0542	0.0115	0.0040	0.0008
150	30	0.0498	0.0080	0.0012	0.0003	1200	60	0.0508	0.0117	0.0045	0.0013
150	36	0.0488	0.0073	0.0020	0.0007	2400	6	0.0407	0.0038	0.0007	0.0000
150	42	0.0475	0.0068	0.0012	0.0002	2400	12	0.0517	0.0068	0.0017	0.0002
150	48	0.0457	0.0078	0.0017	0.0003	2400	18	0.0490	0.0083	0.0017	0.0000
150	54	0.0517	0.0073	0.0022	0.0002	2400	24	0.0485	0.0072	0.0022	0.0007
150	60	0.0548	0.0090	0.0022	0.0008	2400	30	0.0448	0.0073	0.0022	0.0003
300	6	0.0418	0.0050	0.0005	0.0000	2400	36	0.0485	0.0078	0.0020	0.0007
300	12	0.0482	0.0075	0.0005	0.0002	2400	42	0.0522	0.0102	0.0023	0.0007
300	18	0.0568	0.0103	0.0027	0.0010	2400	48	0.0520	0.0102	0.0035	0.0018
300	24	0.0457	0.0093	0.0018	0.0005	2400	54	0.0498	0.0107	0.0032	0.0017
300	30	0.0490	0.0092	0.0020	0.0003	2400	60	0.0500	0.0108	0.0040	0.0022
300	36	0.0465	0.0080	0.0017	0.0005	4800	6	0.0540	0.0043	0.0000	0.0000
300	42	0.0452	0.0082	0.0022	0.0007	4800	12	0.0583	0.0055	0.0008	0.0000
300	48	0.0482	0.0098	0.0030	0.0010	4800	18	0.0583	0.0060	0.0005	0.0002
300	54	0.0498	0.0068	0.0013	0.0000	4800	24	0.0567	0.0092	0.0015	0.0003
300	60	0.0492	0.0085	0.0020	0.0002	4800	30	0.0548	0.0060	0.0012	0.0003
600	6	0.0448	0.0038	0.0003	0.0000	4800	36	0.0467	0.0080	0.0017	0.0007
600	12	0.0507	0.0088	0.0017	0.0002	4800	42	0.0508	0.0097	0.0030	0.0005
600	18	0.0508	0.0088	0.0025	0.0008	4800	48	0.0460	0.0077	0.0025	0.0008
600	24	0.0545	0.0112	0.0030	0.0010	4800	54	0.0432	0.0087	0.0025	0.0008
600	30	0.0508	0.0107	0.0025	0.0007	4800	60	0.0460	0.0095	0.0030	0.0007

Figura 5-11. Correlazione tra autovalori consecutivi (primi 4 autovalori; 70 scenari)

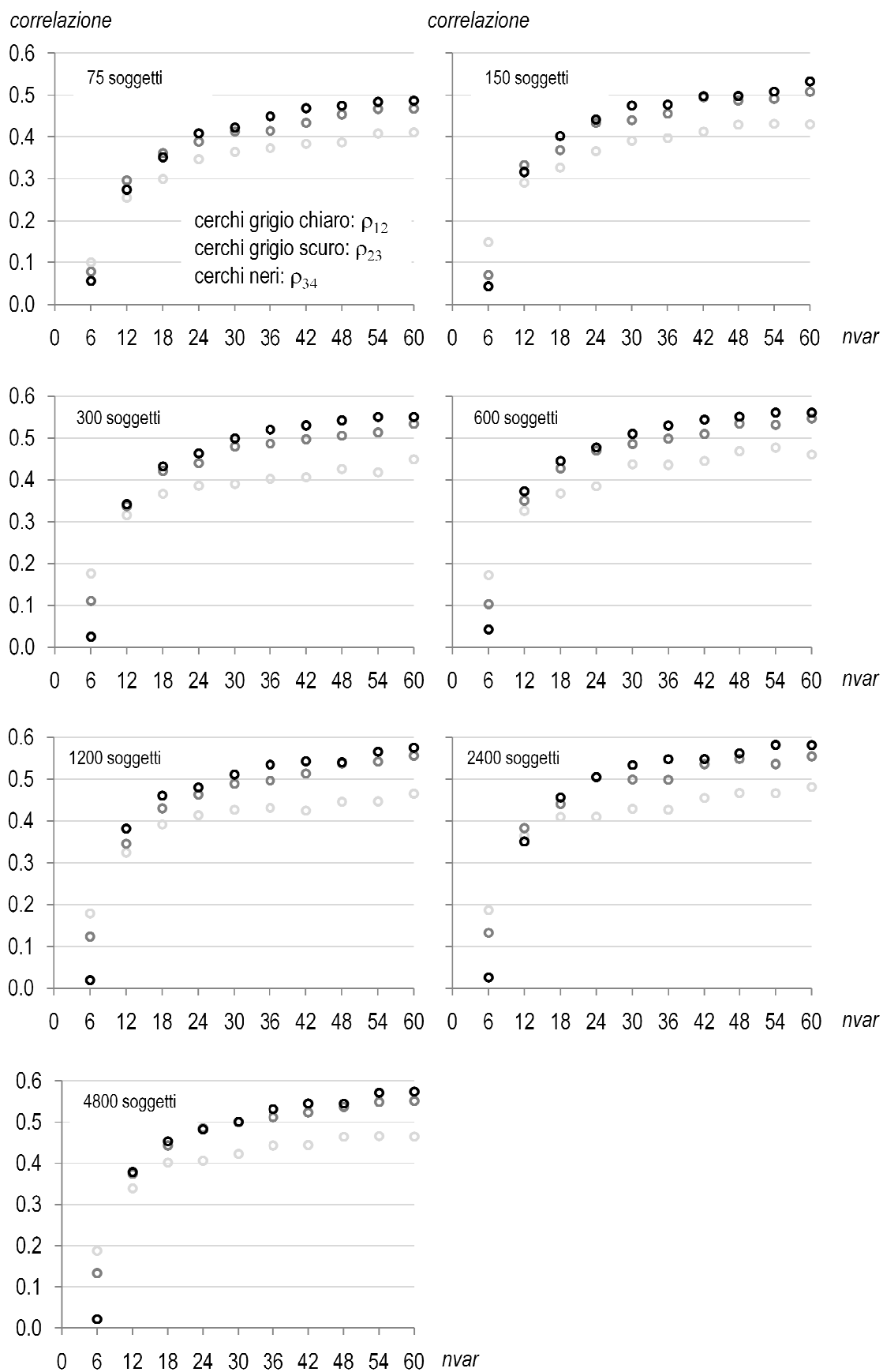


Tabella 5. Differenza tra la probabilità di superare la soglia predetta con il metodo diretto (p_P) e la probabilità di superare la soglia osservata (p_O) da parte del primo autovalore (l_1), o le soglie (t_2, t_3, t_4) da parte degli autovalori successivi (l_2, l_3, l_4), condizionata al superamento della soglia da parte degli autovalori precedenti. Sono riportati solo gli scenari dove tali probabilità mostrano la massima differenza (sottolineata), positiva o negativa, per almeno un autovalore.

scenario		condizionate			
n	k	$p(l_1 > t_1)$	$p(l_2 > t_2 l_1 > t_1)$	$p(l_3 > t_3 l_1 > t_1, l_2 > t_2)$	$p(l_4 > t_4 l_1 > t_1, l_2 > t_2, l_3 > t_3)$
		$\frac{p_P - p_O}{(p_O)}$	$\frac{p_P - p_O}{(p_O)}$	$\frac{p_P - p_O}{(p_O)}$	$\frac{p_P - p_O}{(p_O)}$
150	18	0.0072 (0.0500)	<u>0.0283</u> (0.1467)	-0.0008 (0.1136)	0.0000 (0.2000)
150	54	0.0017 (0.0500)	-0.0181 (0.1600)	-0.0170 (0.3125)	<u>-0.1897</u> (0.2667)
600	24	0.0045 (0.0500)	0.0082 (0.1967)	<u>0.0483</u> (0.2203)	-0.1282 (0.4615)
2400	6	<u>-0.0093</u> (0.0500)	0.0076 (0.0867)	0.0201 (0.1538)	0.0000 (0.0000)
4800	12	<u>0.0083</u> (0.0500)	-0.0257 (0.1200)	-0.0429 (0.1944)	0.0000 (0.0000)
4800	18	<u>0.0083</u> (0.0500)	<u>-0.0438</u> (0.1467)	<u>-0.0758</u> (0.1591)	<u>0.3333</u> (0.0000)

3.2.3 Adeguatezza della funzione

Al fine di valutare l'adeguatezza della funzione per il calcolo del 95° centile è stata usata, per ogni scenario, la differenza assoluta tra il rischio di errore di tipo I relativo alle soglie osservate (p_O) e alle soglie predette con il metodo diretto (p_P). In altri termini, p_O rappresenta la probabilità che l'autovalore superi la soglia osservata e p_P la probabilità che l'autovalore superi la soglia predetta dalla funzione. Nel complesso i risultati evidenziano l'adeguatezza della funzione scelta, presentando una differenza assoluta media soltanto del 3‰ per il primo autovalore, e dello 0.1‰ per il quarto autovalore (Tabella 6).

Tabella 6. Differenza per scenario tra la probabilità di superare la soglia osservata (p_O) e la probabilità di superare la soglia predetta con il metodo diretto (p_P): valori medi e massimi nell'insieme degli scenari

	<i>media di</i> $ p_P - p_O $	<i>max</i> $ p_P - p_O $ (p_O)	<i>scenario</i> <i>corrispondente</i>	
			<i>n</i>	<i>k</i>
<i>almeno il 1° autovalore > 95°</i>	0.003060	0.009340 (0.050000)	2400	6
<i>almeno i primi 2 > 95°</i>	0.000783	0.002665 (0.007333)	150	18
<i>almeno i primi 3 > 95°</i>	0.000233	0.000834 (0.002000)	4800	30
<i>almeno i primi 4 > 95°</i>	0.000102	0.000500 (0.000833)	4800	30

Ovviamente, le differenze per scenario tra le probabilità *condizionate* aumentano per gli autovalori successivi al primo, poiché il numero di campioni su cui sono basate diminuisce e aumentano le fluttuazioni casuali (Tabella 5).

3.3 Calcolo dei valori soglia in funzione della dimensione dello studio e del numero di variabili: metodo indiretto

La distribuzione degli autovalori si scosta dalla distribuzione gaussiana per tutti gli scenari considerati, anche se in misura differente (si vedano le Figure A1-A96 in Appendice). Il secondo metodo per il calcolo di valori soglia corrispondenti al 95° centile è consistito nella normalizzazione della distribuzione dei primi 4 autovalori considerati. Rispetto al metodo diretto, il riferimento ad una normale standard risulta a priori vantaggioso, a parità di performance, poiché a partire dai soli valori di n e k permette il calcolo di qualunque valore soglia (non esclusivamente il 95° centile). Inoltre, il metodo indiretto permette di associare a qualsiasi autovalore osservato un valore di probabilità, sotto ipotesi nulla.

3.3.1 Funzione per il calcolo del 95° centile

Per normalizzare la distribuzione dei primi 4 autovalori, si è fatto ricorso a una trasformazione di Box e Cox a 3 parametri, poiché la trasformazione che tiene conto anche della curtosi, con l'uso della trasformata K (Haynes 1997) non ha migliorato sensibilmente l'approssimazione alla distribuzione normale. Per ogni autovalore e scenario sono stati calcolati i 3 parametri della trasformata di Box e Cox μ (media), ξ (coefficiente di variazione) e λ (parametro di asimmetria):

$$l = \mu(1 + \lambda\xi z)^{\frac{1}{\lambda}},$$

dove l indica l'autovalore e z il rankit corrispondente alla sua posizione nell'insieme ordinato dei 6001 autovalori. Per il calcolo dei rankit si è ricorsi alla PROC RANK, adottando la formula di Blom (Blom, 1958). I valori dei parametri sono stati calcolati facendo uso della PROC NLIN. Entrambe le procedure sono parte del software SAS® (v.9.2, 2008). Ottenuti i parametri si sono ricavati i valori normalizzati (sds_1 - sds_4) dei primi 4 autovalori per ogni scenario:

$$sds = \frac{l - \mu}{\lambda\xi}.$$

Le Tabelle A1-A4 in Appendice mostrano i valori dei parametri μ (media), σ (deviazione standard), γ_1 (asimmetria) e γ_2 (curtosi) delle distribuzioni degli autovalori prima (l_1 - l_4) e dopo normalizzazione (sds_1 -

sds₄). Inoltre, per mostrare l'effetto della normalizzazione, per l_1-l_4 e per sds_1-sds_4 si sono riportati in Appendice i Q-Qplot relativi agli scenari con 6, 24, 42 e 60 variabili per le dimensioni 75, 600 e 4800 soggetti.

Per predire il 95° centile con il metodo indiretto sono state utilizzate tre varianti. Nella variante 1 sono stati sostituiti all'interno della trasformata di Box-Cox i valori calcolati dei parametri (μ , ξ e λ); la variante 2 è equivalente alla variante 1, ma il valore del parametro λ è stato considerato costante al variare del numero di variabili e pari alla media dei 10 valori relativi a ogni dimensione dello studio (μ , ξ e $\bar{\lambda}$). Nella variante 3 i valori del 95° centile sono stati ottenuti sostituendo nella trasformata di Box-Cox i valori $\hat{\mu}$, $\hat{\xi}$ e $\hat{\bar{\lambda}}$, ossia i valori di μ , ξ e $\bar{\lambda}$ predetti da funzioni in n e k (par. 3.3.2).

3.3.2 Interpolazione dei parametri μ , ξ , $\bar{\lambda}$ in funzione di n , k

Per la variante 3, i parametri μ , ξ e $\bar{\lambda}$ sono stati predetti dalle seguenti funzioni³:

$$\mu = 1 + (\alpha n^\beta) k^{\gamma + \delta n^\varepsilon} + \zeta (nk)^\eta$$

$$\xi = (\vartheta + \iota n) k + (\kappa n^\nu) k^{o + \pi \cdot \log(n)}$$

$$\bar{\lambda} = \omega + \rho \cdot \log(n).$$

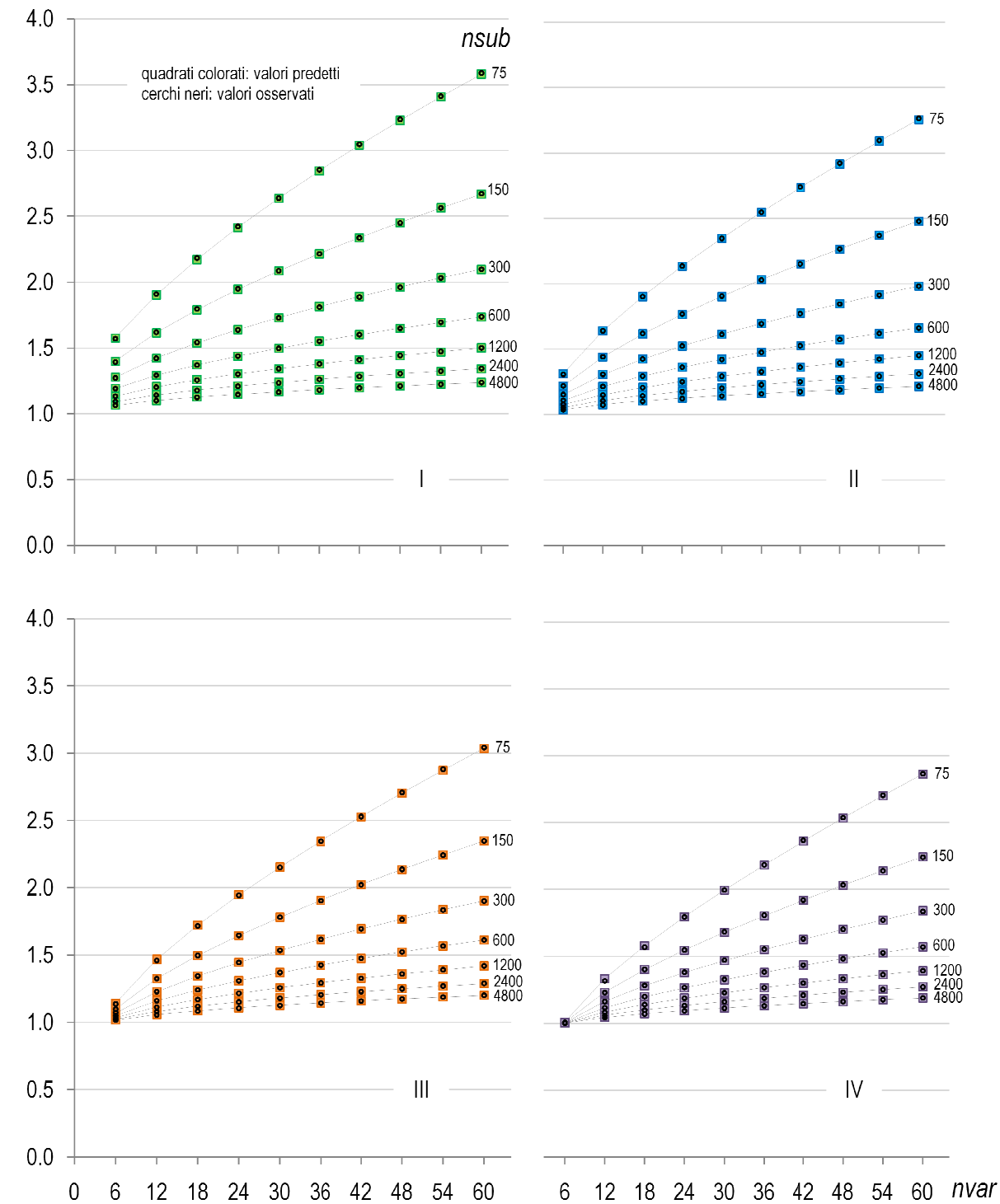
Per ogni autovalore, i quattordici valori per calcolare le soglie sono riportati in Tabella 7. Le Figure 9-12 di seguito riportate evidenziano, nel complesso, l'adeguatezza delle funzioni scelte nel predire il 95° centile.

Tabella 7. Valore dei parametri (metodo indiretto)

Tabella IV. Valore dei parametri (metodo indiretto)															
autovalore	valori numerici dei parametri														
	α	β	γ	δ	ε	ζ	η	θ	ι	κ	ν	o	π	ω	ρ
I	0.725	-0.381	0.347	0.730	-0.174	-0.445	-0.273	-0.000	0.000	-0.417	0.061	1.049	-0.564	8.376	-2.140
II	0.530	-0.372	0.303	0.761	-0.132	-2.395	-0.408	-0.000	0.000	-0.433	0.071	0.888	-0.591	7.239	-1.777
III	0.382	-0.350	0.201	0.884	-0.101	-3.912	-0.429	-0.000	0.000	-0.494	0.085	0.994	-0.639	6.014	-1.432
IV	0.274	-0.322	-0.022	1.136	-0.073	-5.266	-0.436	0.000	0.000	-0.571	0.070	1.464	-0.649	3.444	-0.812

³ Per la sola μ i valori dei parametri sono stati ottenuti con il metodo dei minimi quadrati pesati, usando come peso \sqrt{nk} .

Figura 12-15. 95° centili predetti con il metodo indiretto e 95° centili osservati per ogni scenario (I-IV autovalori)



3.3.3 Rischio d'errore di tipo I e probabilità condizionate

La Tabella 8 riporta i rischi di errore di tipo I (α) calcolati sul complesso dei 70 scenari considerando come soglie i valori osservati del 95° centile della distribuzione degli autovalori e i valori predetti per la stessa soglia, ottenuti per le tre differenti varianti.

Le Tabelle 9, 10 e 11 riportano, per ciascuno dei 70 scenari e per ognuno dei 4 autovalori, le probabilità *congiunte* ottenute utilizzando le soglie osservate e quelle predette con il metodo indiretto. Come per il metodo diretto, anche per il metodo indiretto le probabilità più basse corrispondono agli scenari in cui le correlazioni tra primo e secondo autovalore sono inferiori (in particolare gli scenari a 6 variabili). Nelle Tabelle 12, 13 e 14 sono invece riportate le differenze tra la probabilità *condizionata* di superamento della soglia osservata e la probabilità condizionata di superamento della soglia predetta con il metodo indiretto (varianti 1, 2, 3); queste tabelle contengono solo gli scenari dove tali probabilità mostrano il massimo scostamento positivo o negativo dalla media per almeno un autovalore.

Tabella 8. Frequenza assoluta e percentuale di superamento delle soglie osservate e predette con il metodo indiretto, varianti 1, 2, 3 (media su tutti gli scenari)

	soglie osservate		soglie predette					
			variante 1		variante 2		variante 3	
	freq	%	freq	%	freq	%	freq	%
<i>almeno il 1° autovalore > 95°</i>	21000	5.00	21310	5.07	21266	5.06	21252	5.06
<i>almeno i primi 2 > 95°</i>	3418	0.81	3468	0.83	3474	0.83	3482	0.83
<i>almeno i primi 3 > 95°</i>	800	0.19	804	0.19	806	0.19	805	0.19
<i>almeno i primi 4 > 95°</i>	217	0.05	218	0.05	229	0.05	228	0.05

Tabella 9. Probabilità congiunte calcolate con le soglie predette da μ , ξ e λ per ogni scenario (metodo indiretto, variante 1)

scenario		probabilità congiunte				scenario		probabilità congiunte			
<i>n</i>	<i>k</i>	almeno il 1° autovalore > 95°	almeno i primi 2 > 95°	almeno i primi 3 > 95°	almeno i primi 4 > 95°	<i>n</i>	<i>k</i>	almeno il 1° autovalore > 95°	almeno i primi 2 > 95°	almeno i primi 3 > 95°	almeno i primi 4 > 95°
75	6	0.0498	0.0035	0.0000	0.0000	600	36	0.0528	0.0095	0.0018	0.0007
75	12	0.0485	0.0040	0.0003	0.0000	600	42	0.0502	0.0115	0.0022	0.0005
75	18	0.0507	0.0050	0.0008	0.0000	600	48	0.0503	0.0090	0.0023	0.0010
75	24	0.0497	0.0073	0.0013	0.0002	600	54	0.0542	0.0140	0.0037	0.0022
75	30	0.0473	0.0072	0.0013	0.0005	600	60	0.0523	0.0102	0.0027	0.0010
75	36	0.0495	0.0083	0.0018	0.0007	1200	6	0.0555	0.0052	0.0003	0.0000
75	42	0.0480	0.0082	0.0020	0.0005	1200	12	0.0480	0.0075	0.0005	0.0002
75	48	0.0497	0.0068	0.0012	0.0002	1200	18	0.0487	0.0067	0.0015	0.0003
75	54	0.0497	0.0088	0.0018	0.0000	1200	24	0.0495	0.0088	0.0030	0.0005
75	60	0.0493	0.0095	0.0028	0.0002	1200	30	0.0522	0.0113	0.0022	0.0007
150	6	0.0515	0.0050	0.0000	0.0000	1200	36	0.0532	0.0092	0.0022	0.0003
150	12	0.0518	0.0057	0.0012	0.0000	1200	42	0.0497	0.0077	0.0022	0.0007
150	18	0.0515	0.0082	0.0008	0.0002	1200	48	0.0493	0.0083	0.0023	0.0008
150	24	0.0482	0.0072	0.0015	0.0002	1200	54	0.0527	0.0103	0.0035	0.0007
150	30	0.0502	0.0082	0.0012	0.0003	1200	60	0.0498	0.0103	0.0037	0.0008
150	36	0.0508	0.0087	0.0027	0.0007	2400	6	0.0500	0.0043	0.0007	0.0000
150	42	0.0510	0.0082	0.0012	0.0002	2400	12	0.0532	0.0072	0.0020	0.0002
150	48	0.0490	0.0083	0.0018	0.0005	2400	18	0.0488	0.0085	0.0018	0.0002
150	54	0.0512	0.0073	0.0022	0.0003	2400	24	0.0535	0.0085	0.0027	0.0007
150	60	0.0485	0.0082	0.0020	0.0008	2400	30	0.0495	0.0088	0.0025	0.0007
300	6	0.0512	0.0057	0.0005	0.0000	2400	36	0.0530	0.0090	0.0023	0.0008
300	12	0.0480	0.0067	0.0005	0.0002	2400	42	0.0525	0.0102	0.0023	0.0007
300	18	0.0547	0.0092	0.0025	0.0008	2400	48	0.0508	0.0095	0.0033	0.0017
300	24	0.0472	0.0087	0.0017	0.0005	2400	54	0.0478	0.0092	0.0025	0.0012
300	30	0.0512	0.0093	0.0020	0.0003	2400	60	0.0505	0.0102	0.0037	0.0020
300	36	0.0510	0.0090	0.0018	0.0003	4800	6	0.0512	0.0042	0.0000	0.0000
300	42	0.0490	0.0087	0.0022	0.0007	4800	12	0.0520	0.0060	0.0010	0.0000
300	48	0.0513	0.0105	0.0030	0.0012	4800	18	0.0493	0.0073	0.0013	0.0000
300	54	0.0532	0.0083	0.0022	0.0003	4800	24	0.0535	0.0102	0.0022	0.0005
300	60	0.0520	0.0095	0.0023	0.0003	4800	30	0.0520	0.0078	0.0022	0.0008
600	6	0.0518	0.0043	0.0003	0.0000	4800	36	0.0485	0.0085	0.0020	0.0010
600	12	0.0510	0.0082	0.0013	0.0002	4800	42	0.0533	0.0112	0.0033	0.0008
600	18	0.0508	0.0080	0.0023	0.0007	4800	48	0.0487	0.0085	0.0032	0.0010
600	24	0.0543	0.0103	0.0022	0.0010	4800	54	0.0483	0.0097	0.0028	0.0007
600	30	0.0530	0.0105	0.0022	0.0005	4800	60	0.0513	0.0095	0.0032	0.0008

Tabella 10. Probabilità congiunte calcolate con le soglie predette da μ , ξ e $\bar{\lambda}$ per ogni scenario (metodo indiretto, variante 2)

scenario		probabilità congiunte				scenario		probabilità congiunte			
<i>n</i>	<i>k</i>	<i>almeno il 1° autovalore > 95°</i>	<i>almeno i primi 2 > 95°</i>	<i>almeno i primi 3 > 95°</i>	<i>almeno i primi 4 > 95°</i>	<i>n</i>	<i>k</i>	<i>almeno il 1° autovalore > 95°</i>	<i>almeno i primi 2 > 95°</i>	<i>almeno i primi 3 > 95°</i>	<i>almeno i primi 4 > 95°</i>
75	6	0.0475	0.0032	0.0000	0.0000	600	36	0.0522	0.0090	0.0017	0.0007
75	12	0.0470	0.0038	0.0003	0.0000	600	42	0.0517	0.0117	0.0022	0.0007
75	18	0.0497	0.0048	0.0008	0.0000	600	48	0.0515	0.0092	0.0023	0.0010
75	24	0.0480	0.0068	0.0012	0.0002	600	54	0.0557	0.0143	0.0037	0.0023
75	30	0.0473	0.0072	0.0013	0.0005	600	60	0.0537	0.0105	0.0028	0.0012
75	36	0.0512	0.0083	0.0018	0.0007	1200	6	0.0545	0.0052	0.0003	0.0000
75	42	0.0498	0.0083	0.0020	0.0005	1200	12	0.0433	0.0065	0.0003	0.0002
75	48	0.0502	0.0068	0.0012	0.0002	1200	18	0.0488	0.0068	0.0015	0.0003
75	54	0.0497	0.0088	0.0018	0.0000	1200	24	0.0497	0.0087	0.0030	0.0005
75	60	0.0510	0.0097	0.0028	0.0002	1200	30	0.0522	0.0113	0.0022	0.0007
150	6	0.0452	0.0040	0.0000	0.0000	1200	36	0.0535	0.0092	0.0022	0.0005
150	12	0.0473	0.0052	0.0012	0.0000	1200	42	0.0528	0.0083	0.0025	0.0007
150	18	0.0517	0.0085	0.0010	0.0002	1200	48	0.0522	0.0090	0.0023	0.0008
150	24	0.0475	0.0072	0.0015	0.0002	1200	54	0.0528	0.0107	0.0037	0.0008
150	30	0.0493	0.0077	0.0008	0.0003	1200	60	0.0498	0.0105	0.0037	0.0010
150	36	0.0510	0.0083	0.0025	0.0007	2400	6	0.0547	0.0048	0.0007	0.0000
150	42	0.0537	0.0083	0.0012	0.0002	2400	12	0.0520	0.0070	0.0020	0.0003
150	48	0.0490	0.0083	0.0018	0.0005	2400	18	0.0518	0.0093	0.0020	0.0002
150	54	0.0537	0.0087	0.0025	0.0007	2400	24	0.0523	0.0082	0.0025	0.0007
150	60	0.0527	0.0092	0.0027	0.0012	2400	30	0.0472	0.0087	0.0025	0.0007
300	6	0.0485	0.0057	0.0005	0.0000	2400	36	0.0505	0.0087	0.0020	0.0007
300	12	0.0480	0.0067	0.0005	0.0002	2400	42	0.0520	0.0100	0.0022	0.0007
300	18	0.0523	0.0087	0.0023	0.0008	2400	48	0.0510	0.0095	0.0033	0.0017
300	24	0.0457	0.0078	0.0015	0.0005	2400	54	0.0487	0.0093	0.0025	0.0012
300	30	0.0557	0.0100	0.0018	0.0003	2400	60	0.0497	0.0100	0.0037	0.0020
300	36	0.0512	0.0095	0.0018	0.0003	4800	6	0.0537	0.0043	0.0000	0.0000
300	42	0.0498	0.0087	0.0022	0.0007	4800	12	0.0525	0.0060	0.0010	0.0000
300	48	0.0518	0.0105	0.0030	0.0012	4800	18	0.0495	0.0070	0.0008	0.0000
300	54	0.0533	0.0083	0.0022	0.0003	4800	24	0.0538	0.0103	0.0023	0.0005
300	60	0.0503	0.0098	0.0025	0.0003	4800	30	0.0548	0.0088	0.0027	0.0008
600	6	0.0535	0.0050	0.0003	0.0000	4800	36	0.0475	0.0085	0.0020	0.0010
600	12	0.0473	0.0075	0.0013	0.0002	4800	42	0.0533	0.0113	0.0035	0.0008
600	18	0.0508	0.0080	0.0023	0.0007	4800	48	0.0480	0.0085	0.0033	0.0010
600	24	0.0512	0.0098	0.0022	0.0010	4800	54	0.0455	0.0093	0.0028	0.0008
600	30	0.0510	0.0102	0.0022	0.0005	4800	60	0.0487	0.0092	0.0030	0.0008

Tabella 11. Probabilità congiunte calcolate con le soglie predette da $\hat{\mu}$, $\hat{\xi}$ e $\hat{\lambda}$ per ogni scenario (metodo indiretto, variante 3)

scenario		probabilità congiunte				scenario		probabilità congiunte			
<i>n</i>	<i>k</i>	<i>almeno il 1° autovalore > 95°</i>	<i>almeno i primi 2 > 95°</i>	<i>almeno i primi 3 > 95°</i>	<i>almeno i primi 4 > 95°</i>	<i>n</i>	<i>k</i>	<i>almeno il 1° autovalore > 95°</i>	<i>almeno i primi 2 > 95°</i>	<i>almeno i primi 3 > 95°</i>	<i>almeno i primi 4 > 95°</i>
75	6	0.0463	0.0028	0.0000	0.0000	600	36	0.0477	0.0095	0.0018	0.0007
75	12	0.0560	0.0052	0.0007	0.0000	600	42	0.0493	0.0113	0.0022	0.0007
75	18	0.0583	0.0063	0.0008	0.0000	600	48	0.0508	0.0090	0.0023	0.0010
75	24	0.0548	0.0083	0.0017	0.0002	600	54	0.0567	0.0150	0.0038	0.0023
75	30	0.0507	0.0073	0.0015	0.0007	600	60	0.0570	0.0112	0.0028	0.0010
75	36	0.0567	0.0097	0.0022	0.0008	1200	6	0.0583	0.0062	0.0005	0.0000
75	42	0.0537	0.0097	0.0023	0.0007	1200	12	0.0448	0.0067	0.0003	0.0002
75	48	0.0550	0.0090	0.0020	0.0002	1200	18	0.0463	0.0062	0.0013	0.0003
75	54	0.0538	0.0103	0.0020	0.0000	1200	24	0.0455	0.0078	0.0027	0.0005
75	60	0.0553	0.0122	0.0037	0.0003	1200	30	0.0483	0.0098	0.0017	0.0007
150	6	0.0418	0.0040	0.0000	0.0000	1200	36	0.0512	0.0088	0.0022	0.0007
150	12	0.0565	0.0065	0.0013	0.0000	1200	42	0.0483	0.0072	0.0020	0.0007
150	18	0.0587	0.0100	0.0012	0.0002	1200	48	0.0513	0.0088	0.0025	0.0008
150	24	0.0563	0.0088	0.0018	0.0002	1200	54	0.0592	0.0122	0.0045	0.0008
150	30	0.0523	0.0085	0.0012	0.0003	1200	60	0.0623	0.0147	0.0050	0.0012
150	36	0.0507	0.0080	0.0022	0.0008	2400	6	0.0563	0.0052	0.0008	0.0000
150	42	0.0493	0.0077	0.0012	0.0002	2400	12	0.0478	0.0057	0.0013	0.0000
150	48	0.0473	0.0080	0.0018	0.0005	2400	18	0.0400	0.0055	0.0010	0.0000
150	54	0.0532	0.0077	0.0022	0.0003	2400	24	0.0398	0.0045	0.0010	0.0005
150	60	0.0553	0.0092	0.0023	0.0010	2400	30	0.0367	0.0052	0.0013	0.0003
300	6	0.0428	0.0060	0.0007	0.0000	2400	36	0.0413	0.0057	0.0017	0.0005
300	12	0.0482	0.0073	0.0005	0.0002	2400	42	0.0490	0.0090	0.0020	0.0007
300	18	0.0565	0.0102	0.0025	0.0008	2400	48	0.0530	0.0098	0.0033	0.0017
300	24	0.0457	0.0090	0.0018	0.0005	2400	54	0.0563	0.0125	0.0037	0.0017
300	30	0.0490	0.0092	0.0020	0.0003	2400	60	0.0632	0.0148	0.0055	0.0028
300	36	0.0467	0.0080	0.0017	0.0007	4800	6	0.0737	0.0075	0.0000	0.0000
300	42	0.0453	0.0080	0.0022	0.0010	4800	12	0.0488	0.0043	0.0008	0.0000
300	48	0.0483	0.0098	0.0030	0.0012	4800	18	0.0405	0.0042	0.0003	0.0000
300	54	0.0508	0.0070	0.0013	0.0002	4800	24	0.0392	0.0057	0.0008	0.0000
300	60	0.0503	0.0092	0.0023	0.0002	4800	30	0.0378	0.0038	0.0007	0.0002
600	6	0.0520	0.0048	0.0003	0.0000	4800	36	0.0345	0.0065	0.0010	0.0005
600	12	0.0507	0.0085	0.0015	0.0002	4800	42	0.0453	0.0080	0.0025	0.0005
600	18	0.0488	0.0083	0.0025	0.0008	4800	48	0.0460	0.0080	0.0030	0.0010
600	24	0.0523	0.0100	0.0023	0.0010	4800	54	0.0512	0.0105	0.0030	0.0010
600	30	0.0483	0.0095	0.0022	0.0005	4800	60	0.0663	0.0157	0.0058	0.0013

Tabella 12. Differenza tra la probabilità di superare la soglia predetta con il metodo indiretto, variante 1 (p_P) e la probabilità di superare la soglia osservata (p_O) da parte del primo autovalore (l_1), o le soglie (t_2 , t_3 , t_4) da parte degli autovalori successivi (l_2 , l_3 , l_4), condizionata al superamento della soglia da parte degli autovalori precedenti. Sono riportati solo gli scenari dove tali probabilità mostrano la massima differenza (sottolineata), positiva o negativa, per almeno un autovalore (o per le soglie osservate o per le soglie predette con il metodo diretto o indiretto)

scenario		condizionate			
n	k	$p(l_1 > t_1)$	$p(l_2 > t_2 \mid l_1 > t_1)$	$p(l_3 > t_3 \mid l_1 > t_1, l_2 > t_2)$	$p(l_4 > t_4 \mid l_1 > t_1, l_2 > t_2, l_3 > t_3)$
		$p_P - p_O$ (p_O)	$p_P - p_O$ (p_O)	$p_P - p_O$ (p_O)	$p_P - p_O$ (p_O)
150	18	0.0015 (0.0500)	<u>0.0119</u> (0.1467)	-0.0116 (0.1136)	0.0000 (0.2000)
150	54	0.0012 (0.0500)	<u>-0.0167</u> (0.1600)	-0.0170 (0.3125)	<u>-0.1128</u> (0.2667)
300	24	<u>-0.0028</u> (0.0500)	0.0071 (0.1767)	0.0036 (0.1887)	0.0000 (0.3000)
300	48	0.0013 (0.0500)	0.0012 (0.2033)	-0.0585 (0.3443)	<u>0.1032</u> (0.2857)
600	24	0.0043 (0.0500)	-0.0065 (0.1967)	-0.0107 (0.2203)	0.0000 (0.4615)
1200	6	<u>0.0055</u> (0.0500)	0.0031 (0.0900)	-0.0096 (0.0741)	0.0000 (0.0000)
2400	6	0.0000 (0.0500)	0.0000 (0.0867)	0.0000 (0.1538)	0.0000 (0.0000)
4800	12	0.0020 (0.0500)	-0.0046 (0.1200)	-0.0278 (0.1944)	0.0000 (0.0000)
4800	18	-0.0007 (0.0500)	0.0020 (0.1467)	0.0227 (0.1591)	0.0000 (0.0000)
4800	54	-0.0017 (0.0500)	0.0000 (0.2000)	<u>0.0264</u> (0.2667)	-0.0772 (0.3125)
4800	60	0.0013 (0.0500)	-0.0149 (0.2000)	<u>-0.3167</u> (0.3167)	0.0000 (0.2632)

Tabella 13. Differenza tra la probabilità di superare la soglia predetta con il metodo indiretto, variante 2 (p_P) e la probabilità di superare la soglia osservata (p_O) da parte del primo autovalore (l_1), o le soglie (t_2 , t_3 , t_4) da parte degli autovalori successivi (l_2 , l_3 , l_4), condizionata al superamento della soglia da parte degli autovalori precedenti. Sono riportati solo gli scenari dove tali probabilità mostrano la massima differenza (sottolineata), positiva o negativa, per almeno un autovalore (o per le soglie osservate o per le soglie predette con il metodo diretto o indiretto)

scenario		condizionate			
n	k	$p(l_1 > t_1)$	$p(l_2 > t_2 \mid l_1 > t_1)$	$p(l_3 > t_3 \mid l_1 > t_1, l_2 > t_2)$	$p(l_4 > t_4 \mid l_1 > t_1, l_2 > t_2, l_3 > t_3)$
		$p_P - p_O$ (p_O)	$p_P - p_O$ (p_O)	$p_P - p_O$ (p_O)	$p_P - p_O$ (p_O)
150	18	0.0017 (0.0500)	0.0178 (0.1467)	0.0040 (0.1136)	-0.0333 (0.2000)
150	54	0.0037 (0.0500)	0.0015 (0.1600)	-0.0240 (0.3125)	0.0000 (0.2667)
150	60	0.0027 (0.0500)	0.0107 (0.1633)	<u>0.0460</u> (0.2449)	-0.0625 (0.5000)
300	48	0.0018 (0.0500)	-0.0008 (0.2033)	<u>-0.0585</u> (0.3443)	0.1032 (0.2857)
600	6	0.0035 (0.0500)	<u>0.0201</u> (0.0733)	-0.0242 (0.0909)	0.0000 (0.0000)
600	24	0.0012 (0.0500)	-0.0045 (0.1967)	0.0000 (0.2203)	0.0000 (0.4615)
600	36	0.0022 (0.0500)	<u>-0.0141</u> (0.1867)	-0.0112 (0.1964)	0.0364 (0.3636)
600	54	<u>0.0057</u> (0.0500)	0.0042 (0.2533)	-0.0205 (0.2763)	0.0173 (0.6190)
1200	12	<u>-0.0067</u> (0.0500)	-0.0100 (0.1600)	-0.0112 (0.0625)	<u>0.1667</u> (0.3333)
2400	6	0.0047 (0.0500)	0.0017 (0.0867)	-0.0159 (0.1538)	0.0000 (0.0000)
2400	12	0.0020 (0.0500)	0.0013 (0.1333)	0.0107 (0.2750)	<u>0.1667</u> (0.0000)
4800	12	0.0025 (0.0500)	-0.0057 (0.1200)	-0.0278 (0.1944)	0.0000 (0.0000)
4800	18	-0.0005 (0.0500)	-0.0053 (0.1467)	-0.0400 (0.1591)	0.0000 (0.0000)
4800	30	0.0048 (0.0500)	0.0111 (0.1500)	0.0352 (0.2667)	<u>-0.1042</u> (0.4167)

Tabella 14. Differenza tra la probabilità di superare la soglia predetta con il metodo indiretto, variante 3 (p_P) e la probabilità di superare la soglia osservata (p_O) da parte del primo autovalore (l_1), o le soglie (t_2 , t_3 , t_4) da parte degli autovalori successivi (l_2 , l_3 , l_4), condizionata al superamento della soglia da parte degli autovalori precedenti. Sono riportati solo gli scenari dove tali probabilità mostrano la massima differenza (sottolineata), positiva o negativa, per almeno un autovalore (o per le soglie osservate o per le soglie predette con il metodo diretto o indiretto)

scenario		condizionate			
n	k	$p(l_1 > t_1)$	$p(l_2 > t_2 \mid l_1 > t_1)$	$p(l_3 > t_3 \mid l_1 > t_1, l_2 > t_2)$	$p(l_4 > t_4 \mid l_1 > t_1, l_2 > t_2, l_3 > t_3)$
		$p_P - p_O$ (p_O)	$p_P - p_O$ (p_O)	$p_P - p_O$ (p_O)	$p_P - p_O$ (p_O)
150	18	0.0087 (0.0500)	0.0238 (0.1467)	0.0030 (0.1136)	-0.0571 (0.2000)
150	54	0.0032 (0.0500)	-0.0158 (0.1600)	-0.0299 (0.3125)	-0.1128 (0.2667)
600	24	0.0023 (0.0500)	-0.0056 (0.1967)	0.0130 (0.2203)	-0.0330 (0.4615)
2400	6	0.0063 (0.0500)	0.0050 (0.0867)	0.0074 (0.1538)	0.0000 (0.0000)
2400	24	-0.0102 (0.0500)	-0.0504 (0.1633)	-0.0840 (0.3061)	<u>0.2333</u> (0.2667)
4800	6	<u>0.0237</u> (0.0500)	0.0185 (0.0833)	0.0000 (0.0000)	0.0000 (0.0000)
4800	12	-0.0012 (0.0500)	-0.0313 (0.1200)	-0.0021 (0.1944)	0.0000 (0.0000)
4800	18	-0.0095 (0.0500)	-0.0438 (0.1467)	-0.0791 (0.1591)	0.0000 (0.0000)
4800	24	-0.0108 (0.0500)	<u>-0.0520</u> (0.1967)	-0.0733 (0.2203)	-0.1538 (0.1538)
4800	30	-0.0122 (0.0500)	-0.0487 (0.1500)	<u>-0.0928</u> (0.2667)	<u>-0.1667</u> (0.4167)
4800	36	<u>-0.0155</u> (0.0500)	0.0151 (0.1733)	-0.0769 (0.2308)	0.0000 (0.5000)
4800	60	0.0163 (0.0500)	<u>0.0362</u> (0.2000)	<u>0.0557</u> (0.3167)	-0.0346 (0.2632)

3.3.4 Adeguatezza della funzione

Al fine di valutare l'adeguatezza del metodo indiretto (con particolare riferimento alla variante 3) nell'ottenere il valore predetto del 95° centile è stata usata, per ogni scenario, la differenza assoluta tra il rischio di errore di tipo I relativo alle soglie osservate (p_O) e alle soglie predette ottenute sostituendo $\hat{\mu}$, $\hat{\xi}$ e $\hat{\lambda}$ nella funzione di trasformazione di Box-Cox (p_P). I risultati sono riportati nelle Tabelle 15-17. Con particolare riferimento alla variante 3 (Tabella 17), anche adottando il metodo indiretto i risultati appaiono soddisfacenti: la differenza tra le probabilità varia dal 5‰ per il primo autovalore allo 0.1‰ per il quarto

autovalore. Il metodo indiretto permette inoltre di associare ad un qualsiasi autovalore osservato un valore di probabilità, a partire dai soli parametri noti n e k .

Tabella 15. Differenza per scenario tra la probabilità di superare la soglia osservata (p_o) e la probabilità di superare la soglia predetta calcolata con il metodo indiretto, variante 1 (p_P): valori medi e massimi nell'insieme degli scenari

	<i>media di</i> $ p_P - p_o $	<i>max</i> $ p_P - p_o $ (p_o)	<i>scenario</i> <i>corrispondente</i>	
			n	k
<i>almeno il 1° autovalore > 95°</i>	0.001670	0.005491 (0.050000)	1200	6
<i>almeno i primi 2 > 95°</i>	0.000300	0.001331 (0.012667)	600	54
<i>almeno i primi 3 > 95°</i>	0.000076	0.000500 (0.003500)	300	48
<i>almeno i primi 4 > 95°</i>	0.000036	0.000333 (0.000667)	150	54

Tabella 16. Differenza per scenario tra la probabilità di superare la soglia osservata (p_o) e la probabilità di superare la soglia predetta calcolata con il metodo indiretto, variante 2 (p_P): valori medi e massimi nell'insieme degli scenari

	<i>media di</i> $ p_P - p_o $	<i>max</i> $ p_P - p_o $ (p_o)	<i>scenario</i> <i>corrispondente</i>	
			n	k
<i>almeno il 1° autovalore > 95°</i>	0.002260	0.006674 (0.050000)	1200	12
<i>almeno i primi 2 > 95°</i>	0.000419	0.001664 (0.012667)	600	54
<i>almeno i primi 3 > 95°</i>	0.000119	0.000666 (0.002000)	150 4800	60 30
<i>almeno i primi 4 > 95°</i>	0.000048	0.000333 (0.000000)	2400	12

Tabella 17. Differenza per scenario tra la probabilità di superare la soglia osservata (p_o) e la probabilità di superare la soglia predetta calcolata con il metodo indiretto, variante 3 (p_P): valori medi e massimi nell'insieme degli scenari

	<i>media di</i> $ p_P - p_o $	<i>max</i> $ p_P - p_o $ (p_o)	<i>scenario</i> <i>corrispondente</i>	
			n	k
<i>almeno il 1° autovalore > 95°</i>	0.005308	0.023654 (0.050000)	4800	6
<i>almeno i primi 2 > 95°</i>	0.001395	0.005664 (0.010000)	4800	60
<i>almeno i primi 3 > 95°</i>	0.000398	0.002666 (0.003167)	4800	60
<i>almeno i primi 4 > 95°</i>	0.000121	0.000833 (0.002000)	2400	60

4. Discussione e conclusioni

In questa tesi si è affrontato il problema, cruciale nell'ambito della PCA, del numero di fattori da considerare come espressione di un costrutto latente. Sono stati esplorati 70 diversi scenari, ottenuti combinando differenti valori di n , dimensione dello studio, e k , numero di variabili, e per ciascuno di questi scenari sono stati generati dati per 6001 campioni. Per ogni scenario è stato così possibile ricavare empiricamente la distribuzione degli autovalori e, in particolare, il valore soglia del 95° centile. Questo metodo empirico, denominato Parallel Analysis, fu originariamente proposto da Horn nel 1965.

Come anticipato nel paragrafo 2, Hayton *et al.* (2004) mettono in guardia, nell'uso del metodo della PA, dai rischi dovuti alla presenza di un fattore latente, che potrebbe indurre a ignorare altri fattori latenti, in particolare nel caso in cui n sia piccolo o nel caso in cui un secondo fattore sia basato su pochi item. È stato tuttavia mostrato (Figure 5-11) che il valore (positivo) della correlazione tra le coppie di autovalori consecutivi cresce all'aumentare di k e, anche se in misura minore, di n , ed è plausibile che questo andamento rimanga invariato anche in presenza di uno o più fattori latenti. A differenza dell'articolo di Hayton *et al.*, questa tesi si limita a individuare criteri razionali per la determinazione del numero di componenti da considerare nella PCA, basati sulla distribuzione degli autovalori di matrici di correlazione calcolate su campioni tratti da una distribuzione gaussiana multivariata la cui matrice di correlazione (P) è stata fatta coincidere con la matrice di identità (I) per rappresentare l'assenza di fattori latenti.

Per evitare il ricorso a tabelle estensive di valori soglia per i diversi scenari o la generazione di dati con procedure di calcolo intensivo (quali la PA) si è scelto di esprimere i valori del 95° centile della distribuzione degli autovalori in funzione di n e k . A questo scopo si sono adottati due metodi. Nel metodo diretto si è definita una funzione di n e k in grado di predire il 95° centile della distribuzione osservata dei primi 4 autovalori. Nel metodo indiretto, dopo aver normalizzato le distribuzioni dei primi 4 autovalori con la trasformazione di Box-Cox a tre parametri (media μ , coefficiente di variazione ξ , parametro di asimmetria λ), si sono definite funzioni in n e k in grado di predire i valori di tali parametri, valori poi utilizzati per predire il 95° centile.

Nell'uno e nell'altro caso le funzioni approssimano il 95° centile osservato in un numero di scenari relativamente limitato, né si può escludere che si sarebbe potuta ottenere un'approssimazione anche migliore estendendo il range di n e k esplorato e infittendo la griglia (ad esempio incrementando k di 3 anziché di 6, ed n di un fattore 1,41 anziché di 2). Il numero di campioni generati in questo studio (6001), inoltre, è sufficientemente grande per ottenere stime precise del rischio d'errore di tipo I per il primo autovalore; l'affidabilità di queste stime è però più bassa per il secondo e, a maggior ragione, per il terzo e il quarto autovalore. Tuttavia i risultati appaiono soddisfacenti e di pratica utilità: in entrambi i casi, come mostrato nelle Tabelle 6 e 17, le funzioni usate per la predizione del 95° centile sono risultate adeguate. In particolare, per il primo autovalore la media delle differenze in valore assoluto tra le

probabilità è del 3‰ adottando il metodo diretto e del 5‰ con il metodo indiretto. Il secondo metodo permette inoltre di associare a qualsiasi autovalore osservato un valore di probabilità, sotto ipotesi nulla. Ulteriori studi in questo campo dovrebbero concentrarsi su come e quanto la distribuzione degli autovalori di matrici di correlazione campionarie dipenda dalla forma della distribuzione genitrice (sia essa asimmetrica, leptocurtica, plurimodale), e sulla possibilità di estensione delle funzioni predittive al caso in cui esistano fattori latenti, attraverso l'aggiunta di un parametro che tien conto della varianza spiegata da tali fattori.

5. Appendice

Tabella A1. Valori dei parametri μ , σ , γ_1 γ_2 e modulo dei rapporti tra asimmetria e curtosi (l_1 e sds_1)

scenario		media		ds		asimmetria		$\frac{l_1 \text{ asim}}{sds_1 \text{ asim}}$	curtosi		$\frac{l_1 \text{ curt}}{sds_1 \text{ curt}}$
n	k	l_1	sds_1	l_1	sds_1	l_1	sds_1		l_1	sds_1	
75	6	1.3991	0.0000	0.0973	0.9990	0.5479	0.0160	34.18	0.5417	-0.0721	7.51
75	12	1.7196	0.0001	0.1094	0.9985	0.5040	0.0259	19.48	0.3760	-0.0742	5.07
75	18	1.9782	0.0000	0.1157	0.9992	0.4644	0.0129	35.98	0.3763	-0.0571	6.59
75	24	2.2103	0.0000	0.1218	0.9992	0.4315	0.0128	33.60	0.2997	-0.0488	6.14
75	30	2.4223	0.0000	0.1258	0.9998	0.4475	-0.0016	282.84	0.4067	0.0067	61.16
75	36	2.6225	0.0000	0.1327	0.9998	0.4691	0.0017	279.24	0.4401	-0.0168	26.18
75	42	2.8115	0.0000	0.1363	0.9996	0.4340	0.0041	104.92	0.3673	-0.0333	11.02
75	48	2.9948	0.0000	0.1394	1.0004	0.3980	-0.0189	21.03	0.3934	0.0922	4.27
75	54	3.1693	0.0000	0.1418	0.9992	0.3686	0.0203	18.19	0.1633	-0.1241	1.32
75	60	3.3385	0.0000	0.1449	0.9994	0.4057	0.0104	39.06	0.2865	-0.0695	4.12
150	6	1.2772	0.0000	0.0653	0.9982	0.4392	0.0490	8.96	0.1689	-0.2381	0.71
150	12	1.4918	0.0000	0.0723	0.9987	0.4317	0.0288	14.97	0.2311	-0.1234	1.87
150	18	1.6630	0.0000	0.0760	0.9995	0.5031	0.0044	115.14	0.5295	-0.0050	106.75
150	24	1.8135	0.0000	0.0785	0.9997	0.4355	0.0030	144.82	0.3979	-0.0435	9.14
150	30	1.9486	0.0000	0.0791	0.9999	0.3992	-0.0036	112.45	0.3571	-0.0149	24.05
150	36	2.0733	0.0000	0.0807	1.0001	0.4234	-0.0077	55.16	0.4157	0.0059	70.10
150	42	2.1920	0.0000	0.0821	1.0001	0.4430	-0.0101	43.85	0.4562	0.0225	20.25
150	48	2.3052	0.0000	0.0835	0.9995	0.3848	0.0060	63.94	0.2720	-0.0459	5.93
150	54	2.4140	0.0000	0.0857	0.9989	0.4218	0.0215	19.65	0.2483	-0.0844	2.94
150	60	2.5188	0.0000	0.0879	0.9986	0.4124	0.0356	11.58	0.1620	-0.1457	1.11
300	6	1.1931	0.0000	0.0448	0.9990	0.4420	0.0221	20.05	0.3106	-0.1046	2.97
300	12	1.3385	0.0000	0.0478	0.9994	0.4738	0.0096	49.37	0.4418	-0.0511	8.64
300	18	1.4536	0.0000	0.0493	0.9986	0.3999	0.0346	11.57	0.1448	-0.1378	1.05
300	24	1.5513	0.0000	0.0499	1.0001	0.3899	-0.0061	63.56	0.3887	0.0140	27.86
300	30	1.6394	0.0000	0.0508	1.0000	0.4758	-0.0060	79.42	0.5188	0.0001	4323.58
300	36	1.7211	0.0000	0.0511	1.0001	0.3971	-0.0102	38.97	0.3617	0.0510	7.09
300	42	1.7977	0.0000	0.0517	0.9998	0.3893	-0.0031	127.43	0.3305	0.0000	-
300	48	1.8697	0.0000	0.0527	0.9991	0.3797	0.0184	20.64	0.1836	-0.0683	2.69
300	54	1.9387	0.0000	0.0531	0.9992	0.3635	0.0172	21.09	0.1626	-0.0670	2.43
300	60	2.0055	0.0000	0.0533	0.9991	0.3084	0.0284	10.86	0.0073	-0.1414	0.05
600	6	1.1357	0.0001	0.0315	0.9976	0.5389	0.0462	11.67	0.4229	-0.1359	3.11
600	12	1.2359	0.0000	0.0325	0.9991	0.3759	0.0248	15.16	0.1618	-0.1313	1.23
600	18	1.3130	0.0000	0.0326	0.9995	0.4516	0.0077	58.44	0.4064	-0.0674	6.03
600	24	1.3795	0.0000	0.0332	0.9992	0.3810	0.0193	19.73	0.1954	-0.0951	2.06
600	30	1.4378	0.0000	0.0334	0.9995	0.3877	0.0047	83.12	0.2607	0.0087	29.97
600	36	1.4917	0.0000	0.0333	0.9996	0.3795	0.0035	108.72	0.2640	-0.0002	1199.95
600	42	1.5416	0.0000	0.0341	0.9993	0.4385	0.0141	31.18	0.3406	-0.0686	4.97
600	48	1.5889	0.0000	0.0343	0.9994	0.4221	0.0058	72.54	0.3232	0.0106	30.60
600	54	1.6340	0.0000	0.0344	0.9992	0.4057	0.0155	26.20	0.2527	-0.0338	7.49
600	60	1.6769	0.0000	0.0345	0.9992	0.3955	0.0156	25.32	0.2221	-0.0300	7.42
1200	6	1.0949	0.0000	0.0220	0.9982	0.4489	0.0460	9.77	0.2090	-0.1756	1.19
1200	12	1.1648	0.0000	0.0220	0.9992	0.3176	0.0230	13.79	0.0600	-0.1226	0.49
1200	18	1.2178	0.0000	0.0224	0.9994	0.4327	0.0063	68.39	0.3553	-0.0123	28.84
1200	24	1.2626	0.0000	0.0224	0.9994	0.4128	0.0066	62.21	0.3167	-0.0225	14.05
1200	30	1.3025	0.0000	0.0225	0.9992	0.3942	0.0200	19.74	0.2299	-0.0831	2.77
1200	36	1.3387	0.0000	0.0226	0.9993	0.3975	0.0136	29.15	0.2434	-0.0210	11.59
1200	42	1.3722	0.0000	0.0223	0.9994	0.4232	0.0080	52.71	0.3410	-0.0395	8.64
1200	48	1.4040	0.0000	0.0222	0.9995	0.4109	0.0083	49.54	0.3125	-0.0440	7.11
1200	54	1.4340	0.0000	0.0223	0.9992	0.3594	0.0216	16.68	0.1650	-0.1194	1.38
1200	60	1.4623	0.0000	0.0222	0.9992	0.3460	0.0211	16.43	0.1334	-0.1208	1.10
2400	6	1.0664	0.0000	0.0152	1.0001	0.4920	-0.0031	158.39	0.5720	-0.0371	15.43
2400	12	1.1150	0.0000	0.0154	0.9992	0.3437	0.0225	15.27	0.1176	-0.1180	1.00
2400	18	1.1517	0.0000	0.0153	1.0002	0.4076	-0.0131	31.16	0.4169	0.0535	7.79
2400	24	1.1827	0.0000	0.0149	0.9998	0.3162	-0.0019	163.07	0.1906	0.0379	5.04
2400	30	1.2097	0.0000	0.0149	0.9993	0.2777	0.0207	13.39	0.0144	-0.1400	0.10
2400	36	1.2348	0.0000	0.0149	0.9994	0.2796	0.0186	14.99	0.0334	-0.1244	0.27
2400	42	1.2578	0.0000	0.0150	0.9992	0.3100	0.0217	14.30	0.0571	-0.1003	0.57
2400	48	1.2792	0.0000	0.0151	0.9998	0.3189	-0.0015	209.11	0.2287	-0.0132	17.31
2400	54	1.2993	0.0000	0.0151	0.9999	0.3266	-0.0037	87.69	0.2563	-0.0065	39.25
2400	60	1.3184	0.0000	0.0147	0.9997	0.2716	0.0012	234.12	0.1424	-0.0227	6.27
4800	6	1.0474	0.0000	0.0107	0.9989	0.4159	0.0293	14.20	0.2372	-0.1312	1.81
4800	12	1.0810	0.0000	0.0108	0.9992	0.3920	0.0161	24.30	0.2397	-0.0689	3.48
4800	18	1.1065	0.0000	0.0108	0.9994	0.3650	0.0140	25.99	0.2107	-0.0656	3.21
4800	24	1.1280	0.0000	0.0104	0.9991	0.3612	0.0224	16.13	0.1396	-0.0469	2.98
4800	30	1.1468	0.0000	0.0103	1.0001	0.3845	-0.0100	38.42	0.3459	0.0614	5.64
4800	36	1.1637	0.0000	0.0101	0.9998	0.3089	0.0020	155.09	0.2124	-0.0424	5.01
4800	42	1.1793	0.0000	0.0100	0.9993	0.3218	0.0169	19.01	0.1054	-0.0699	1.51
4800	48	1.1938	0.0000	0.0100	0.9996	0.2980	0.0103	29.01	0.1350	-0.0770	1.75
4800	54	1.2077	0.0000	0.0098	0.9997	0.2563	0.0050	51.33	0.0962	-0.0398	2.42
4800	60	1.2208	0.0000	0.0098	0.9996	0.2436	0.0087	27.98	0.0509	-0.0772	0.66

Tabella A2. Valori dei parametri μ , σ , γ_1 γ_2 e modulo dei rapporti tra asimmetria e curtosi (l_2 e sds_2)

scenario		media		ds		asimmetria		$\frac{l_2 \text{ asimm}}{sds_2 \text{ asimm}}$	curtosi		$\frac{l_2 \text{ curt}}{sds_2 \text{ curt}}$
n	k	l_2	sds_2	l_2	sds_2	l_2	sds_2		l_2	sds_2	
75	6	1.1969	0.0000	0.0638	0.9994	0.2563	0.0195	13.14	-0.0347	-0.1468	0.24
75	12	1.5093	0.0000	0.0764	0.9993	0.3211	0.0156	20.63	0.0884	-0.0806	1.10
75	18	1.7636	0.0000	0.0820	0.9999	0.2998	-0.0019	154.16	0.2251	-0.0253	8.91
75	24	1.9871	0.0000	0.0869	1.0000	0.2750	-0.0081	33.82	0.1700	0.0714	2.38
75	30	2.1919	0.0000	0.0903	0.9999	0.2950	-0.0046	64.12	0.1957	0.0240	8.16
75	36	2.3859	0.0000	0.0948	1.0001	0.2816	-0.0120	23.47	0.2365	0.0657	3.60
75	42	2.5710	0.0000	0.0986	0.9996	0.2408	0.0049	49.65	0.0574	-0.0240	2.39
75	48	2.7471	0.0000	0.1014	0.9999	0.2617	-0.0046	56.46	0.1388	0.0422	3.29
75	54	2.9158	0.0000	0.1056	0.9995	0.2863	0.0158	18.14	0.0334	-0.0891	0.38
75	60	3.0824	0.0000	0.1088	0.9996	0.2357	0.0141	16.67	-0.0054	-0.1319	0.04
150	6	1.1409	0.0000	0.0452	0.9988	0.3326	0.0415	8.01	-0.0363	-0.2137	0.17
150	12	1.3542	0.0000	0.0512	0.9993	0.2646	0.0315	8.39	-0.1033	-0.2261	0.46
150	18	1.5245	0.0000	0.0552	0.9996	0.3399	0.0063	53.58	0.2261	-0.0391	5.78
150	24	1.6722	0.0000	0.0568	0.9999	0.2526	-0.0081	31.24	0.1676	0.0587	2.85
150	30	1.8059	0.0000	0.0588	0.9997	0.2612	0.0017	155.28	0.1156	-0.0232	4.98
150	36	1.9291	0.0000	0.0590	0.9996	0.2083	0.0091	22.77	-0.0130	-0.0696	0.19
150	42	2.0462	0.0000	0.0614	0.9994	0.2270	0.0192	11.84	-0.0656	-0.1494	0.44
150	48	2.1570	0.0000	0.0635	0.9995	0.2656	0.0162	16.43	0.0108	-0.1045	0.10
150	54	2.2629	0.0000	0.0641	0.9998	0.3391	0.0000	7372.17	0.2446	-0.0138	17.79
150	60	2.3657	0.0000	0.0653	0.9996	0.2658	0.0072	36.81	0.0868	-0.0529	1.64
300	6	1.0997	0.0000	0.0317	0.9994	0.3279	0.0126	26.07	0.1384	-0.0767	1.80
300	12	1.2478	0.0000	0.0349	0.9996	0.2596	0.0037	69.64	0.0929	-0.0291	3.19
300	18	1.3628	0.0000	0.0369	0.9998	0.2436	0.0023	105.25	0.0838	-0.0308	2.72
300	24	1.4616	0.0000	0.0373	0.9998	0.2235	0.0001	4216.42	0.0981	-0.0234	4.20
300	30	1.5492	0.0000	0.0378	0.9998	0.2032	-0.0022	90.44	0.0946	0.0043	21.91
300	36	1.6301	0.0000	0.0386	0.9997	0.2587	0.0002	1226.21	0.1283	-0.0097	13.24
300	42	1.7059	0.0000	0.0393	0.9997	0.2194	0.0077	28.57	0.0291	-0.0747	0.39
300	48	1.7778	0.0000	0.0398	0.9999	0.2048	-0.0056	36.64	0.1211	0.0498	2.43
300	54	1.8464	0.0000	0.0403	0.9998	0.2093	-0.0024	87.18	0.1047	0.0044	23.57
300	60	1.9121	0.0000	0.0411	0.9996	0.2286	0.0042	54.74	0.0652	-0.0457	1.43
600	6	1.0708	0.0000	0.0220	0.9989	0.3252	0.0397	8.20	-0.0238	-0.2061	0.12
600	12	1.1740	0.0000	0.0239	1.0004	0.3168	-0.0214	14.79	0.3213	0.1131	2.84
600	18	1.2526	0.0000	0.0249	0.9998	0.2674	-0.0007	406.40	0.1449	-0.0084	17.29
600	24	1.3197	0.0000	0.0253	0.9998	0.2707	-0.0030	90.44	0.1799	0.0005	391.17
600	30	1.3790	0.0000	0.0254	0.9997	0.2337	0.0032	73.87	0.0756	-0.0238	3.18
600	36	1.4325	0.0000	0.0258	0.9997	0.2192	0.0028	78.29	0.0625	-0.0325	1.92
600	42	1.4830	0.0000	0.0257	0.9998	0.2387	0.0040	59.91	0.1363	-0.0058	23.54
600	48	1.5304	0.0000	0.0263	1.0003	0.2322	-0.0187	12.45	0.2898	0.1328	2.18
600	54	1.5755	0.0000	0.0263	0.9995	0.2520	0.0122	20.62	0.0401	-0.0712	0.56
600	60	1.6181	0.0000	0.0260	1.0001	0.2250	-0.0140	16.10	0.2388	0.0970	2.46
1200	6	1.0503	0.0000	0.0156	0.9990	0.3220	0.0328	9.80	0.0186	-0.1601	0.12
1200	12	1.1223	0.0000	0.0167	0.9998	0.2607	-0.0048	54.47	0.1651	0.0252	6.55
1200	18	1.1770	0.0000	0.0169	0.9996	0.2566	0.0107	24.08	0.0452	-0.0673	0.67
1200	24	1.2228	0.0000	0.0171	0.9997	0.2219	0.0046	48.22	0.0612	-0.0671	0.91
1200	30	1.2628	0.0000	0.0173	0.9996	0.2410	0.0084	28.73	0.0423	-0.0625	0.68
1200	36	1.2993	0.0000	0.0171	0.9997	0.2263	0.0025	91.40	0.0811	-0.0330	2.46
1200	42	1.3332	0.0000	0.0170	0.9997	0.2533	0.0031	81.41	0.0974	-0.0091	10.73
1200	48	1.3653	0.0000	0.0169	0.9997	0.2437	0.0031	79.36	0.0943	-0.0317	2.98
1200	54	1.3952	0.0000	0.0170	0.9996	0.2464	0.0093	26.56	0.0446	-0.0685	0.65
1200	60	1.4235	0.0000	0.0171	0.9995	0.2353	0.0217	10.84	-0.0631	-0.1737	0.36
2400	6	1.0356	0.0000	0.0109	0.9994	0.2724	0.0139	19.61	0.0484	-0.0734	0.66
2400	12	1.0861	0.0000	0.0116	0.9994	0.2829	0.0197	14.39	0.0327	-0.1227	0.27
2400	18	1.1238	0.0000	0.0119	1.0000	0.2446	-0.0065	37.37	0.1998	0.0280	7.13
2400	24	1.1553	0.0000	0.0118	0.9997	0.1931	0.0054	35.62	0.0221	-0.0604	0.37
2400	30	1.1832	0.0000	0.0116	0.9999	0.2108	-0.0087	24.14	0.1542	0.0698	2.21
2400	36	1.2084	0.0000	0.0115	0.9997	0.1598	0.0075	21.41	-0.0384	-0.0857	0.45
2400	42	1.2318	0.0000	0.0116	0.9998	0.1971	-0.0020	96.51	0.0930	0.0017	53.77
2400	48	1.2534	0.0000	0.0116	0.9997	0.1807	0.0024	74.04	0.0340	-0.0345	0.99
2400	54	1.2737	0.0000	0.0115	0.9998	0.1887	-0.0004	454.60	0.0642	-0.0006	103.56
2400	60	1.2930	0.0000	0.0115	0.9996	0.1876	0.0069	27.08	0.0006	-0.0648	0.01
4800	6	1.0252	0.0000	0.0078	0.9992	0.2906	0.0336	8.65	-0.0467	-0.1991	0.23
4800	12	1.0606	0.0000	0.0081	0.9996	0.2176	0.0245	8.86	-0.1164	-0.2106	0.55
4800	18	1.0873	0.0000	0.0082	1.0000	0.2024	-0.0128	15.80	0.2064	0.0975	2.12
4800	24	1.1093	0.0000	0.0080	0.9997	0.2202	0.0003	683.85	0.0829	0.0047	17.48
4800	30	1.1286	0.0000	0.0079	0.9998	0.2202	-0.0016	135.15	0.0904	0.0266	3.40
4800	36	1.1459	0.0000	0.0080	0.9998	0.1867	-0.0066	28.17	0.1261	0.0587	2.15
4800	42	1.1618	0.0000	0.0079	0.9997	0.2095	0.0074	28.40	0.0371	-0.0746	0.50
4800	48	1.1766	0.0000	0.0078	0.9996	0.2189	0.0103	21.31	0.0105	-0.1063	0.10
4800	54	1.1905	0.0000	0.0078	0.9996	0.1970	0.0098	20.14	-0.0135	-0.0941	0.14
4800	60	1.2038	0.0000	0.0078	0.9997	0.1395	0.0081	17.30	-0.0754	-0.1154	0.65

Tabella A3. Valori dei parametri μ , σ , γ_1 γ_2 e modulo dei rapporti tra asimmetria e curtosi (l_3 e sds_3)

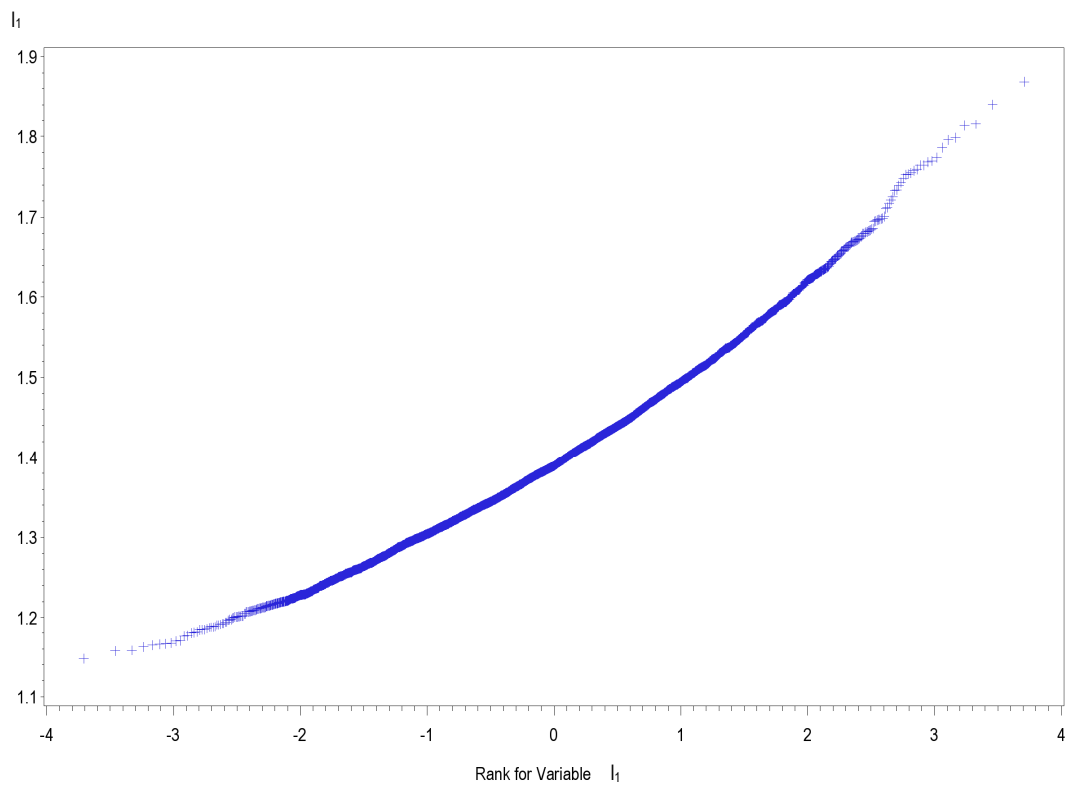
scenario		media		ds		asimmetria		$\frac{l_3 \text{ asimm}}{sds_3 \text{ asimm}}$	curtosi		$\frac{l_3 \text{ curt}}{sds_3 \text{ curt}}$
n	k	l_3	sds_3	l_3	sds_3	l_3	sds_3		l_3	sds_3	
75	6	1.0509	0.0000	0.0511	0.9999	0.0851	-0.0027	31.69	0.0445	0.0532	0.84
75	12	1.3568	0.0000	0.0622	0.9997	0.1709	0.0096	17.81	-0.0697	-0.1215	0.57
75	18	1.6054	0.0000	0.0674	0.9997	0.1612	0.0060	26.84	-0.0228	-0.0824	0.28
75	24	1.8247	0.0000	0.0718	0.9998	0.1953	-0.0001	3255.50	0.0969	-0.0202	4.79
75	30	2.0241	0.0000	0.0758	0.9997	0.1527	0.0043	35.37	-0.0258	-0.0782	0.33
75	36	2.2124	0.0000	0.0792	0.9997	0.1949	0.0007	289.99	0.0560	-0.0101	5.55
75	42	2.3921	0.0000	0.0806	0.9997	0.1864	0.0037	49.73	0.0314	-0.0447	0.70
75	48	2.5651	0.0000	0.0861	0.9996	0.1818	0.0062	29.57	-0.0111	-0.0835	0.13
75	54	2.7302	0.0000	0.0882	0.9997	0.2505	0.0030	83.00	0.0943	-0.0345	2.73
75	60	2.8918	0.0000	0.0897	0.9997	0.2028	0.0027	74.66	0.0494	-0.0413	1.20
150	6	1.0386	0.0000	0.0358	1.0001	0.1650	-0.0131	12.64	0.1438	0.1323	1.09
150	12	1.2527	0.0000	0.0429	0.9997	0.2354	0.0047	49.85	0.0811	-0.0652	1.24
150	18	1.4212	0.0000	0.0450	0.9996	0.2031	0.0091	22.34	-0.0089	-0.1003	0.09
150	24	1.5674	0.0000	0.0478	0.9997	0.1573	0.0068	23.07	-0.0427	-0.0874	0.49
150	30	1.6998	0.0000	0.0489	0.9997	0.1593	0.0043	37.47	-0.0039	-0.0643	0.06
150	36	1.8207	0.0000	0.0499	0.9998	0.1858	-0.0030	61.44	0.0879	0.0248	3.55
150	42	1.9361	0.0000	0.0520	0.9998	0.2038	-0.0013	154.25	0.0871	0.0002	435.45
150	48	2.0473	0.0000	0.0532	0.9999	0.2009	-0.0053	37.69	0.1062	0.0538	1.97
150	54	2.1521	0.0000	0.0538	0.9997	0.2072	0.0009	230.42	0.0640	-0.0108	5.92
150	60	2.2538	0.0000	0.0560	0.9996	0.2470	0.0041	59.73	0.0771	-0.0329	2.34
300	6	1.0287	0.0000	0.0252	0.9999	0.2118	-0.0041	51.79	0.0963	0.0453	2.12
300	12	1.1784	0.0000	0.0294	0.9996	0.1819	0.0174	10.47	-0.1252	-0.1705	0.73
300	18	1.2946	0.0000	0.0308	0.9998	0.1470	-0.0025	58.44	0.0575	0.0080	7.16
300	24	1.3934	0.0000	0.0315	0.9998	0.1434	0.0000	4096.86	0.0665	-0.0008	86.39
300	30	1.4814	0.0000	0.0327	0.9998	0.1145	0.0046	24.94	-0.0643	-0.0984	0.65
300	36	1.5621	0.0000	0.0334	0.9999	0.1701	-0.0055	31.21	0.1012	0.0356	2.84
300	42	1.6378	0.0000	0.0339	0.9999	0.2078	-0.0050	41.26	0.1273	0.0423	3.01
300	48	1.7091	0.0000	0.0340	0.9997	0.1987	0.0026	75.83	0.0512	-0.0404	1.27
300	54	1.7773	0.0000	0.0347	0.9998	0.1954	-0.0026	74.77	0.0919	0.0148	6.20
300	60	1.8424	0.0000	0.0356	0.9998	0.2268	-0.0010	222.38	0.0947	0.0110	8.60
600	6	1.0211	0.0000	0.0179	0.9998	0.1165	-0.0026	45.15	0.0579	0.0344	1.68
600	12	1.1270	0.0000	0.0206	0.9998	0.1569	0.0036	43.62	-0.0041	-0.0531	0.08
600	18	1.2070	0.0000	0.0211	0.9998	0.1922	-0.0026	74.82	0.0820	0.0242	3.40
600	24	1.2744	0.0000	0.0215	0.9998	0.1662	0.0004	474.91	0.0519	-0.0227	2.29
600	30	1.3340	0.0000	0.0216	0.9998	0.2315	-0.0014	163.06	0.1173	0.0020	57.48
600	36	1.3883	0.0000	0.0222	0.9997	0.2150	0.0017	124.99	0.0748	-0.0147	5.10
600	42	1.4383	0.0000	0.0222	0.9997	0.2067	-0.0004	519.30	0.0807	-0.0032	25.37
600	48	1.4860	0.0000	0.0222	0.9998	0.1363	0.0111	12.28	-0.1294	-0.1734	0.75
600	54	1.5312	0.0000	0.0225	0.9997	0.1755	0.0004	398.03	0.0506	-0.0199	2.55
600	60	1.5744	0.0000	0.0230	0.9997	0.2000	0.0004	502.49	0.0647	-0.0053	12.12
1200	6	1.0150	0.0000	0.0125	1.0003	0.2093	-0.0191	10.97	0.2194	0.1566	1.40
1200	12	1.0897	0.0000	0.0142	0.9997	0.1969	0.0078	25.27	-0.0119	-0.0715	0.17
1200	18	1.1454	0.0000	0.0147	0.9996	0.1627	0.0128	12.76	-0.0928	-0.1371	0.68
1200	24	1.1921	0.0000	0.0147	0.9997	0.1743	0.0043	40.63	0.0028	-0.0512	0.05
1200	30	1.2329	0.0000	0.0146	0.9997	0.1659	0.0029	56.66	0.0141	-0.0282	0.50
1200	36	1.2697	0.0000	0.0145	0.9998	0.1369	-0.0034	39.80	0.0868	0.0416	2.09
1200	42	1.3038	0.0000	0.0146	0.9999	0.1743	-0.0111	15.75	0.1814	0.1052	1.72
1200	48	1.3363	0.0000	0.0147	0.9997	0.1537	0.0047	33.05	-0.0138	-0.0407	0.34
1200	54	1.3665	0.0000	0.0150	0.9998	0.1913	-0.0026	73.19	0.0918	0.0098	9.35
1200	60	1.3950	0.0000	0.0151	0.9997	0.1737	0.0048	36.45	-0.0044	-0.0572	0.08
2400	6	1.0109	0.0000	0.0090	1.0002	0.2370	-0.0123	19.21	0.2341	0.1093	2.14
2400	12	1.0636	0.0000	0.0100	0.9997	0.1733	0.0077	22.47	-0.0263	-0.0657	0.40
2400	18	1.1023	0.0000	0.0101	0.9997	0.1099	0.0068	16.15	-0.0974	-0.1183	0.82
2400	24	1.1346	0.0000	0.0102	0.9999	0.1735	-0.0086	20.21	0.1490	0.0815	1.83
2400	30	1.1629	0.0000	0.0102	0.9998	0.1458	0.0019	78.38	0.0386	-0.0252	1.53
2400	36	1.1884	0.0000	0.0101	0.9998	0.1497	-0.0037	40.82	0.0869	0.0361	2.41
2400	42	1.2121	0.0000	0.0101	0.9998	0.1457	0.0101	14.47	-0.0930	-0.1175	0.79
2400	48	1.2338	0.0000	0.0101	0.9997	0.1384	0.0024	58.38	-0.0015	-0.0401	0.04
2400	54	1.2544	0.0000	0.0101	0.9999	0.1143	-0.0062	18.37	0.1109	0.0861	1.29
2400	60	1.2737	0.0000	0.0101	0.9997	0.1205	-0.0005	252.70	0.0357	-0.0020	18.12
4800	6	1.0078	0.0000	0.0064	0.9998	0.2430	-0.0045	53.97	0.1531	0.0202	7.58
4800	12	1.0449	0.0000	0.0070	0.9998	0.1520	-0.0011	142.82	0.0598	-0.0115	5.20
4800	18	1.0724	0.0000	0.0070	0.9998	0.0934	0.0084	11.09	-0.1327	-0.1397	0.95
4800	24	1.0949	0.0000	0.0069	0.9999	0.1532	-0.0058	26.58	0.1244	0.0471	2.64
4800	30	1.1145	0.0000	0.0069	0.9997	0.1457	0.0038	38.31	-0.0124	-0.0530	0.23
4800	36	1.1321	0.0000	0.0069	0.9997	0.1573	0.0112	14.00	-0.0807	-0.1217	0.66
4800	42	1.1484	0.0000	0.0068	0.9999	0.1050	-0.0044	23.81	0.1009	0.0699	1.44
4800	48	1.1633	0.0000	0.0069	0.9997	0.1695	0.0083	20.33	-0.0375	-0.0865	0.43
4800	54	1.1774	0.0000	0.0068	0.9997	0.1512	0.0024	62.44	0.0092	-0.0383	0.24
4800	60	1.1908	0.0000	0.0068	0.9998	0.0899	0.0057	15.90	-0.1017	-0.1159	0.88

Tabella A4. Valori dei parametri μ , σ , γ_1 γ_2 e modulo dei rapporti tra asimmetria e curtosi (l_4 e sds_4)

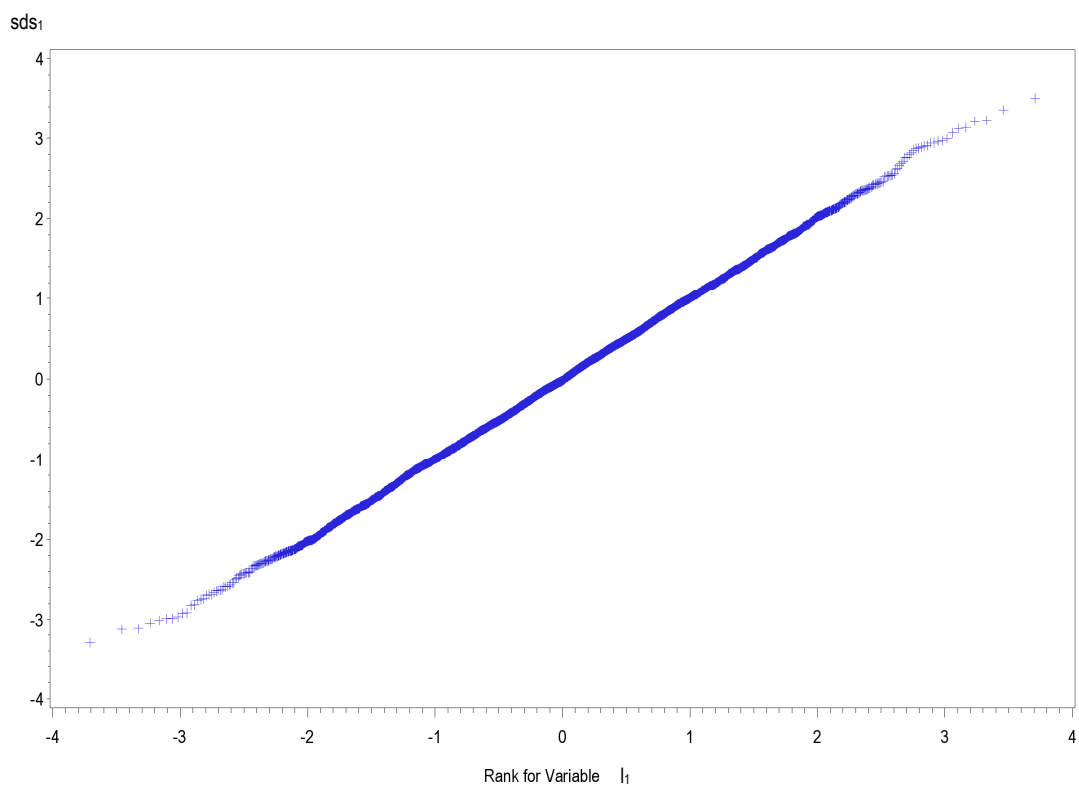
scenario		media		ds		asimmetria		$\frac{l_4 \text{ asimm}}{sds_4 \text{ asimm}}$	curtosi		$\frac{l_4 \text{ curt}}{sds_4 \text{ curt}}$
n	k	l_4	sds_4	l_4	sds_4	l_4	sds_4		l_4	sds_4	
75	6	0.9206	0.0000	0.0522	0.9997	-0.2538	-0.0253	10.04	-0.0199	-0.1499	0.13
75	12	1.2267	0.0000	0.0537	0.9997	0.1092	0.0000	2873.95	0.0145	-0.0026	5.54
75	18	1.4704	0.0000	0.0589	0.9997	0.2245	0.0020	114.84	0.0772	-0.0281	2.75
75	24	1.6865	0.0000	0.0640	0.9997	0.1511	0.0001	1034.66	0.0258	-0.0055	4.68
75	30	1.8829	0.0000	0.0673	0.9998	0.1196	0.0082	14.60	-0.1084	-0.1143	0.95
75	36	2.0667	0.0000	0.0698	0.9998	0.1385	0.0029	47.25	-0.0203	-0.0328	0.62
75	42	2.2430	0.0000	0.0732	0.9998	0.1560	-0.0035	45.19	0.0719	0.0316	2.28
75	48	2.4114	0.0000	0.0753	0.9997	0.1344	0.0093	14.45	-0.1111	-0.1429	0.78
75	54	2.5739	0.0000	0.0777	0.9996	0.1809	0.0075	24.10	-0.0263	-0.0837	0.31
75	60	2.7335	0.0000	0.0795	0.9999	0.1407	-0.0039	35.74	0.0936	0.0351	2.67
150	6	0.9475	0.0000	0.0371	0.9997	-0.2448	-0.0073	33.49	0.1015	-0.0372	2.73
150	12	1.1664	0.0000	0.0377	0.9997	0.1469	0.0118	12.44	-0.1151	-0.1463	0.79
150	18	1.3345	0.0000	0.0407	0.9997	0.1535	0.0071	21.64	-0.0511	-0.0861	0.59
150	24	1.4790	0.0000	0.0419	0.9998	0.1578	-0.0019	81.45	0.0742	0.0049	15.03
150	30	1.6103	0.0000	0.0434	0.9997	0.1762	0.0091	19.33	-0.0475	-0.0827	0.57
150	36	1.7312	0.0000	0.0448	0.9997	0.1948	0.0045	43.34	0.0267	-0.0550	0.48
150	42	1.8450	0.0000	0.0454	0.9999	0.1432	-0.0075	18.96	0.1280	0.0958	1.34
150	48	1.9539	0.0000	0.0477	0.9999	0.1864	-0.0063	29.45	0.1332	0.0483	2.76
150	54	2.0577	0.0000	0.0485	0.9997	0.1346	0.0019	69.80	0.0014	-0.0288	0.05
150	60	2.1585	0.0000	0.0496	0.9998	0.1050	0.0036	28.92	-0.0495	-0.0628	0.79
300	6	0.9643	0.0000	0.0259	0.9998	-0.2747	-0.0014	202.70	0.1524	0.0200	7.61
300	12	1.1188	0.0000	0.0261	0.9997	0.1620	0.0066	24.56	-0.0314	-0.0729	0.43
300	18	1.2368	0.0000	0.0274	0.9997	0.1664	0.0060	27.63	-0.0179	-0.0743	0.24
300	24	1.3353	0.0000	0.0283	0.9997	0.1645	0.0020	81.46	0.0224	-0.0359	0.62
300	30	1.4230	0.0000	0.0292	0.9997	0.1556	0.0064	24.43	-0.0438	-0.0677	0.65
300	36	1.5042	0.0000	0.0298	0.9997	0.1484	0.0057	26.15	-0.0410	-0.0668	0.61
300	42	1.5799	0.0000	0.0304	0.9998	0.0978	-0.0011	88.46	0.0204	0.0171	1.20
300	48	1.6509	0.0000	0.0305	0.9997	0.1432	0.0037	38.53	-0.0091	-0.0512	0.18
300	54	1.7192	0.0000	0.0312	0.9997	0.1650	0.0043	38.22	0.0045	-0.0574	0.08
300	60	1.7833	0.0000	0.0313	0.9997	0.1266	0.0030	42.58	-0.0214	-0.0587	0.36
600	6	0.9757	0.0000	0.0178	0.9996	-0.2557	-0.0185	13.80	0.0218	-0.1457	0.15
600	12	1.0857	0.0000	0.0181	0.9997	0.1781	0.0082	21.61	-0.0287	-0.0900	0.32
600	18	1.1677	0.0000	0.0189	0.9997	0.1581	0.0028	56.69	0.0052	-0.0312	0.17
600	24	1.2354	0.0000	0.0192	0.9998	0.1688	-0.0004	419.88	0.0448	0.0110	4.06
600	30	1.2958	0.0000	0.0197	0.9998	0.1540	-0.0031	48.92	0.0837	0.0210	3.98
600	36	1.3502	0.0000	0.0197	0.9997	0.1783	0.0016	111.30	0.0427	-0.0246	1.74
600	42	1.4008	0.0000	0.0200	0.9997	0.1549	0.0056	27.62	-0.0303	-0.0710	0.43
600	48	1.4485	0.0000	0.0201	0.9997	0.1653	0.0130	12.69	-0.0863	-0.1237	0.70
600	54	1.4931	0.0000	0.0202	0.9998	0.1684	-0.0024	69.59	0.0856	0.0139	6.15
600	60	1.5364	0.0000	0.0203	0.9998	0.1871	-0.0015	126.83	0.0846	-0.0122	6.91
1200	6	0.9831	0.0000	0.0128	0.9997	-0.2314	-0.0083	27.82	0.0429	-0.0725	0.59
1200	12	1.0613	0.0000	0.0128	0.9997	0.1358	0.0036	38.18	-0.0205	-0.0505	0.41
1200	18	1.1182	0.0000	0.0131	0.9998	0.1275	-0.0012	107.59	0.0242	0.0060	4.03
1200	24	1.1654	0.0000	0.0130	0.9998	0.1412	-0.0016	87.83	0.0548	0.0066	8.36
1200	30	1.2071	0.0000	0.0133	0.9999	0.1362	-0.0078	17.36	0.1242	0.0907	1.37
1200	36	1.2447	0.0000	0.0134	0.9997	0.1314	0.0001	1662.91	0.0210	-0.0119	1.77
1200	42	1.2786	0.0000	0.0135	0.9998	0.1398	-0.0011	128.97	0.0492	-0.0023	21.30
1200	48	1.3110	0.0000	0.0134	1.0000	0.1604	-0.0010	153.92	0.1827	0.0636	2.87
1200	54	1.3414	0.0000	0.0135	0.9999	0.1898	-0.0016	118.74	0.1437	0.0235	6.12
1200	60	1.3703	0.0000	0.0137	0.9999	0.1610	-0.0004	447.17	0.1116	0.0119	9.36
2400	6	0.9883	0.0000	0.0088	1.0001	-0.2241	0.0128	17.53	0.1765	0.1082	1.63
2400	12	1.0435	0.0000	0.0089	0.9998	0.1647	-0.0012	140.73	0.0603	0.0053	11.32
2400	18	1.0837	0.0000	0.0091	0.9997	0.1512	0.0041	37.32	-0.0110	-0.0527	0.21
2400	24	1.1167	0.0000	0.0092	0.9998	0.1078	0.0039	27.46	-0.0500	-0.0584	0.85
2400	30	1.1454	0.0000	0.0091	0.9998	0.1053	-0.0019	54.45	0.0347	0.0231	1.50
2400	36	1.1711	0.0000	0.0092	0.9997	0.1226	0.0048	25.39	-0.0501	-0.0830	0.60
2400	42	1.1950	0.0000	0.0092	0.9998	0.0902	0.0050	18.12	-0.0969	-0.1146	0.85
2400	48	1.2170	0.0000	0.0092	0.9997	0.1270	0.0001	1587.88	0.0245	-0.0022	11.35
2400	54	1.2378	0.0000	0.0092	0.9998	0.1231	-0.0007	179.68	0.0200	0.0013	15.88
2400	60	1.2573	0.0000	0.0092	0.9997	0.0941	0.0030	31.04	-0.0490	-0.0626	0.78
4800	6	0.9919	0.0000	0.0064	0.9998	-0.2825	0.0005	535.04	0.1553	0.0383	4.06
4800	12	1.0311	0.0000	0.0064	0.9998	0.1313	-0.0003	497.39	0.0544	0.0104	5.23
4800	18	1.0592	0.0000	0.0064	0.9997	0.1266	0.0091	13.98	-0.1030	-0.1340	0.77
4800	24	1.0823	0.0000	0.0064	0.9997	0.0999	0.0041	24.33	-0.0471	-0.0578	0.81
4800	30	1.1023	0.0000	0.0063	0.9999	0.0868	-0.0028	30.99	0.0558	0.0267	2.09
4800	36	1.1202	0.0000	0.0062	0.9998	0.0957	0.0017	57.50	-0.0206	-0.0306	0.68
4800	42	1.1366	0.0000	0.0063	0.9997	0.1429	0.0067	21.18	-0.0557	-0.1138	0.49
4800	48	1.1518	0.0000	0.0062	0.9997	0.1302	0.0103	12.63	-0.0985	-0.1250	0.79
4800	54	1.1661	0.0000	0.0062	0.9998	0.0939	0.0018	51.01	-0.0275	-0.0578	0.48
4800	60	1.1797	0.0000	0.0062	0.9997	0.0900	0.0038	23.63	-0.0760	-0.0966	0.79

Figura A1-A24. Q-Qplot per l_1 - l_4 e sds_1 - sds_4 relativi agli scenari con 6 variabili per 75, 600 e 4800 soggetti

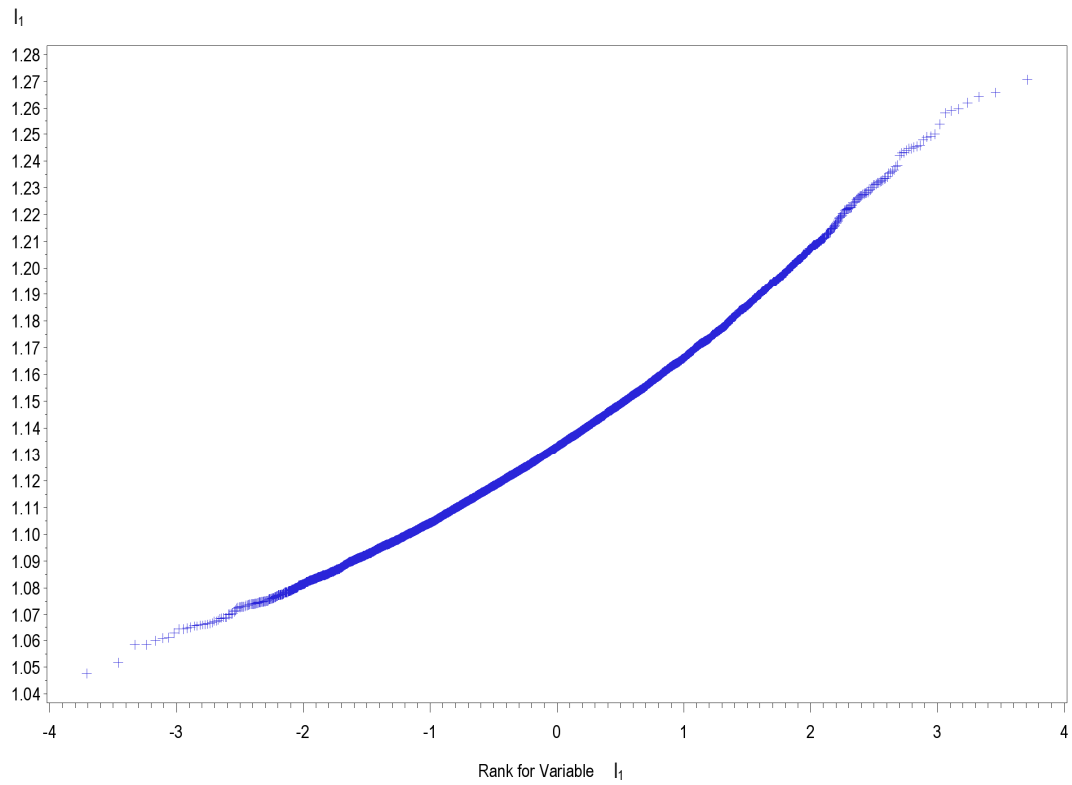
nsub=75 nvar=6



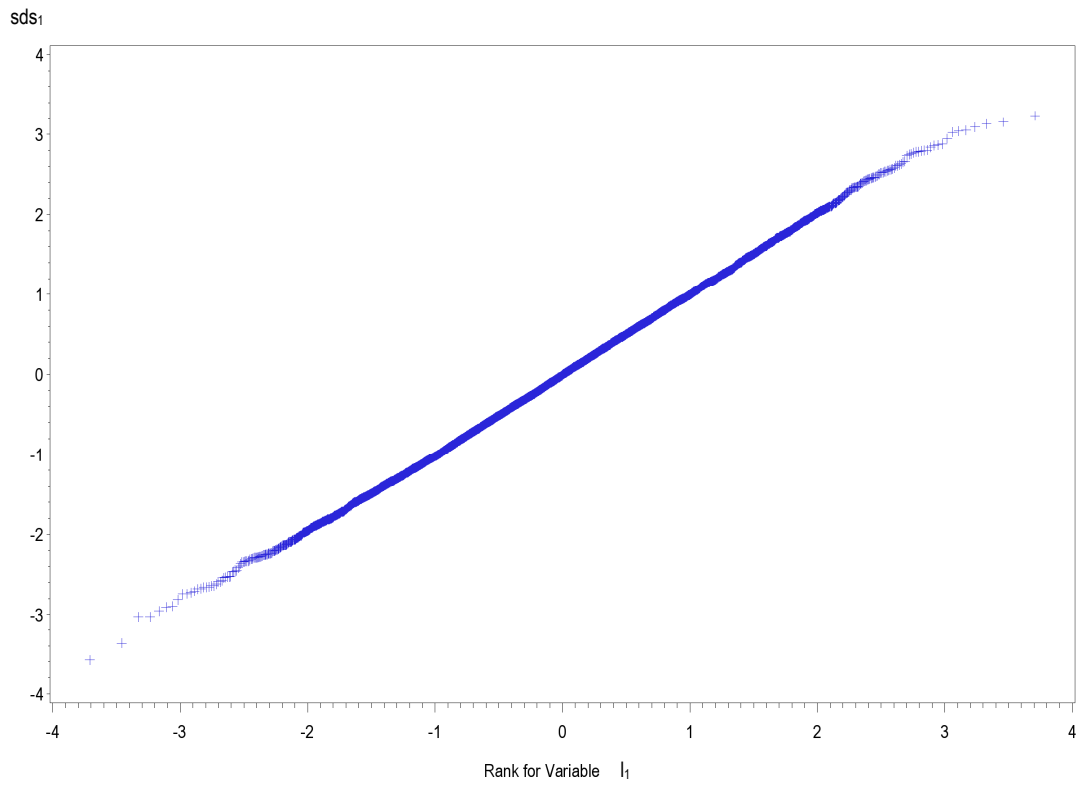
nsub=75 nvar=6



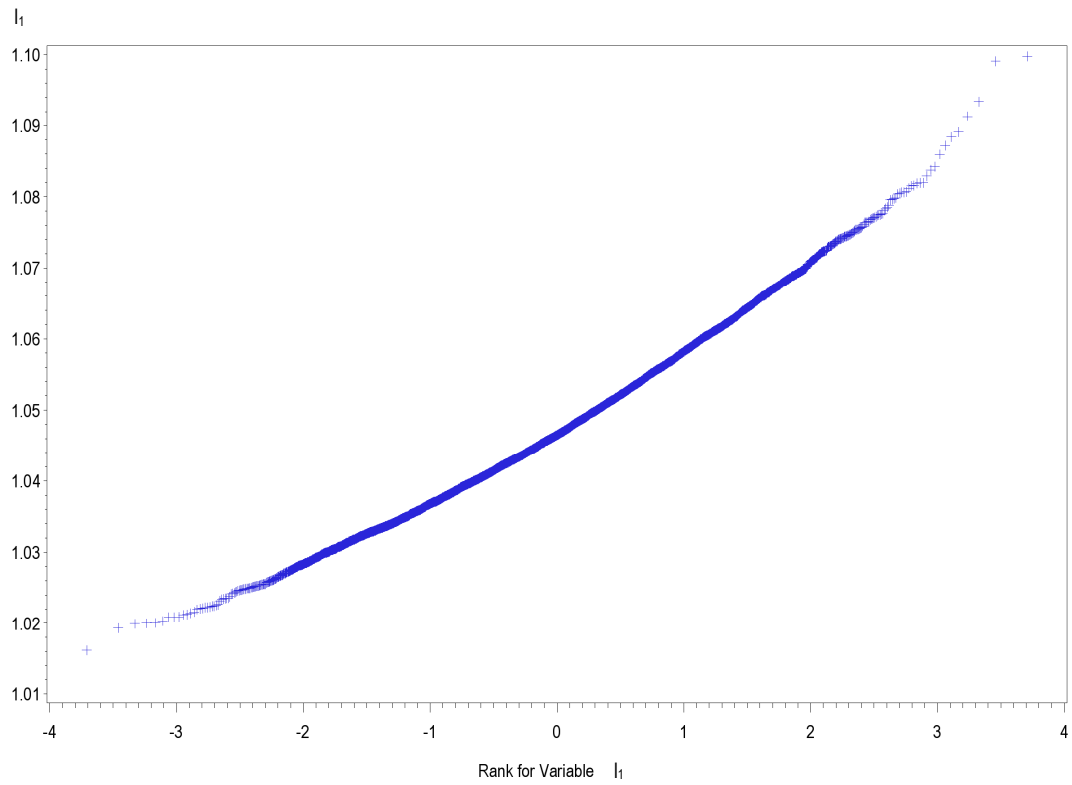
nsub=600 nvar=6



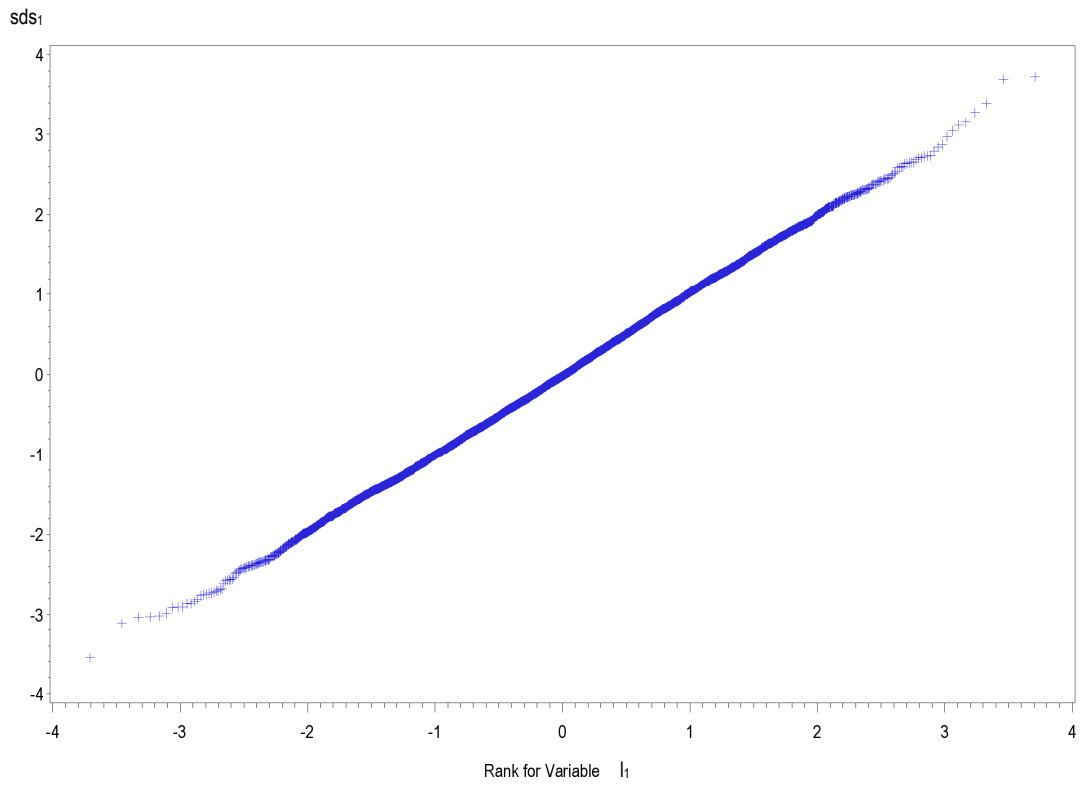
nsub=600 nvar=6



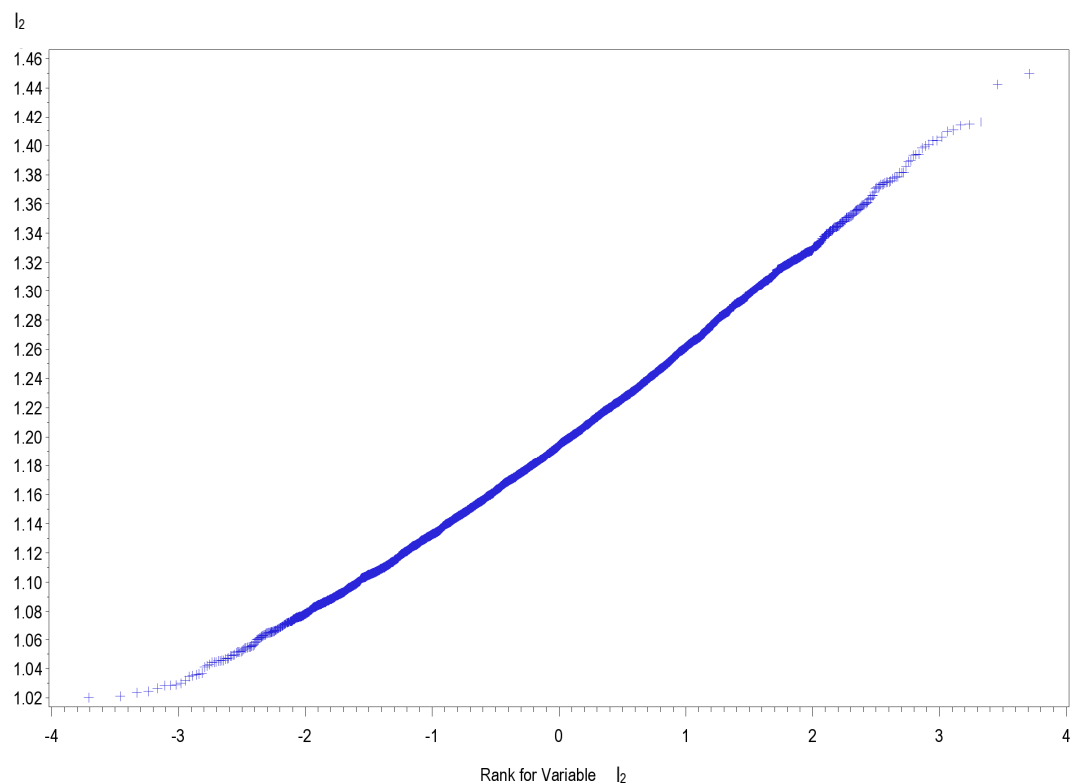
nsub=4800 nvar=6



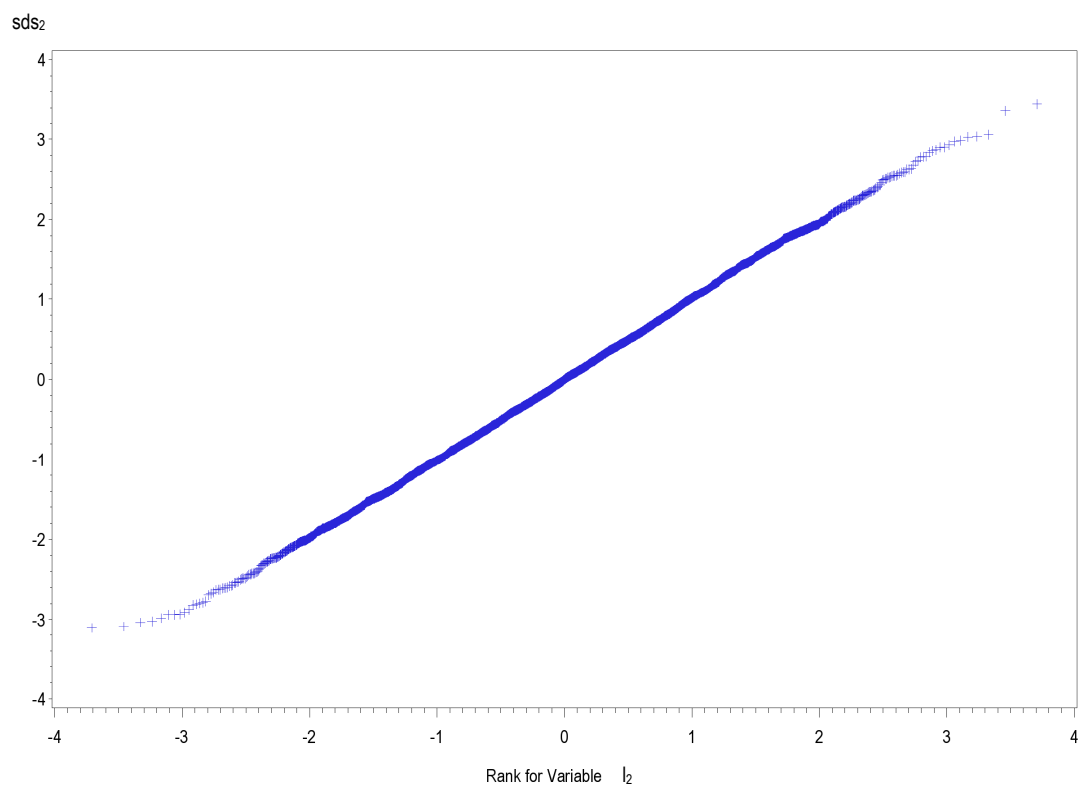
nsub=4800 nvar=6



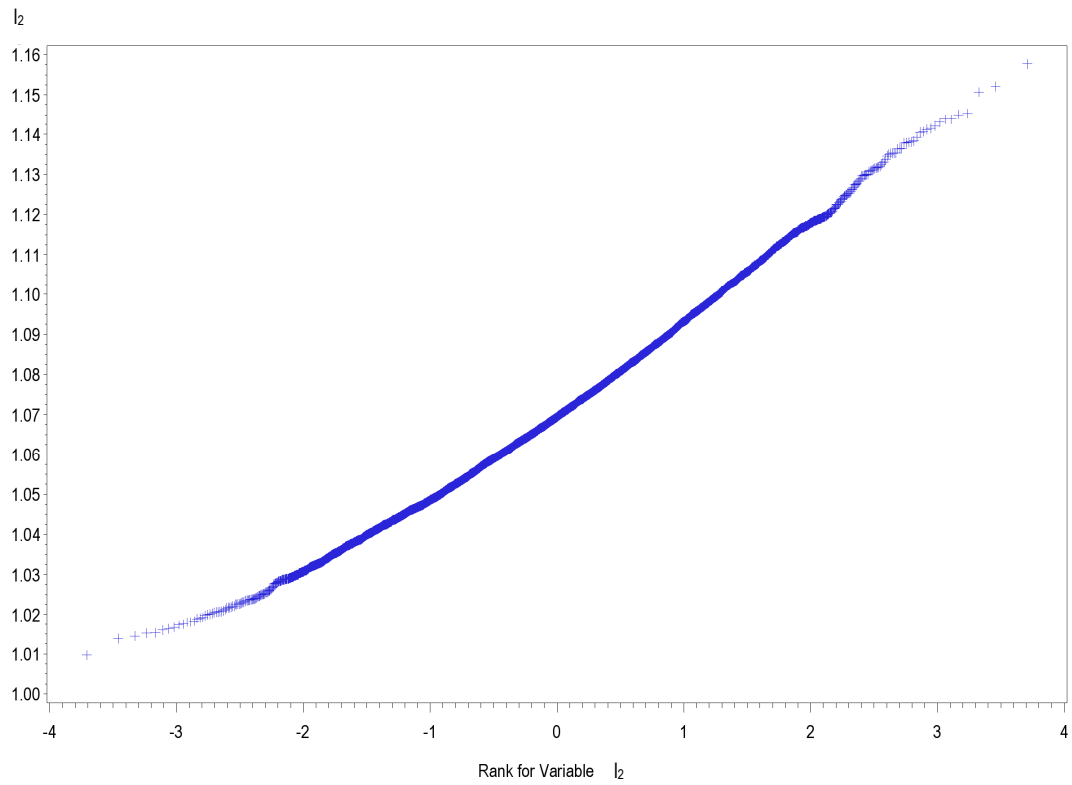
nsub=75 nvar=6



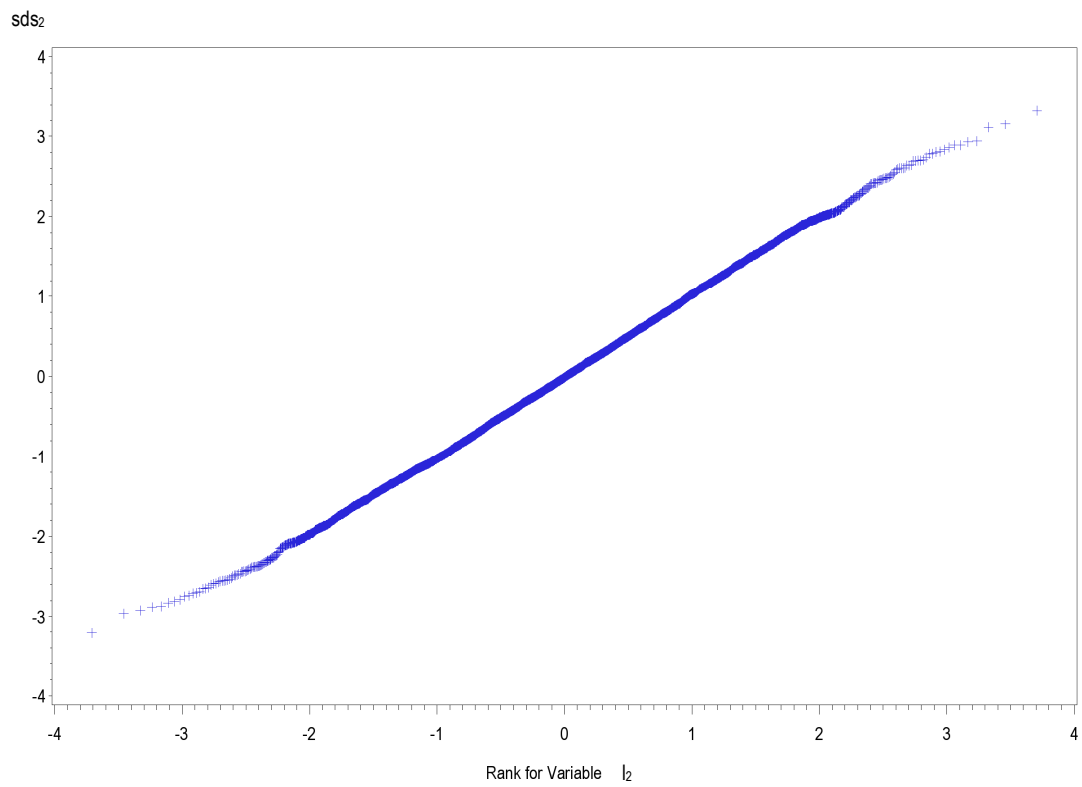
nsub=75 nvar=6



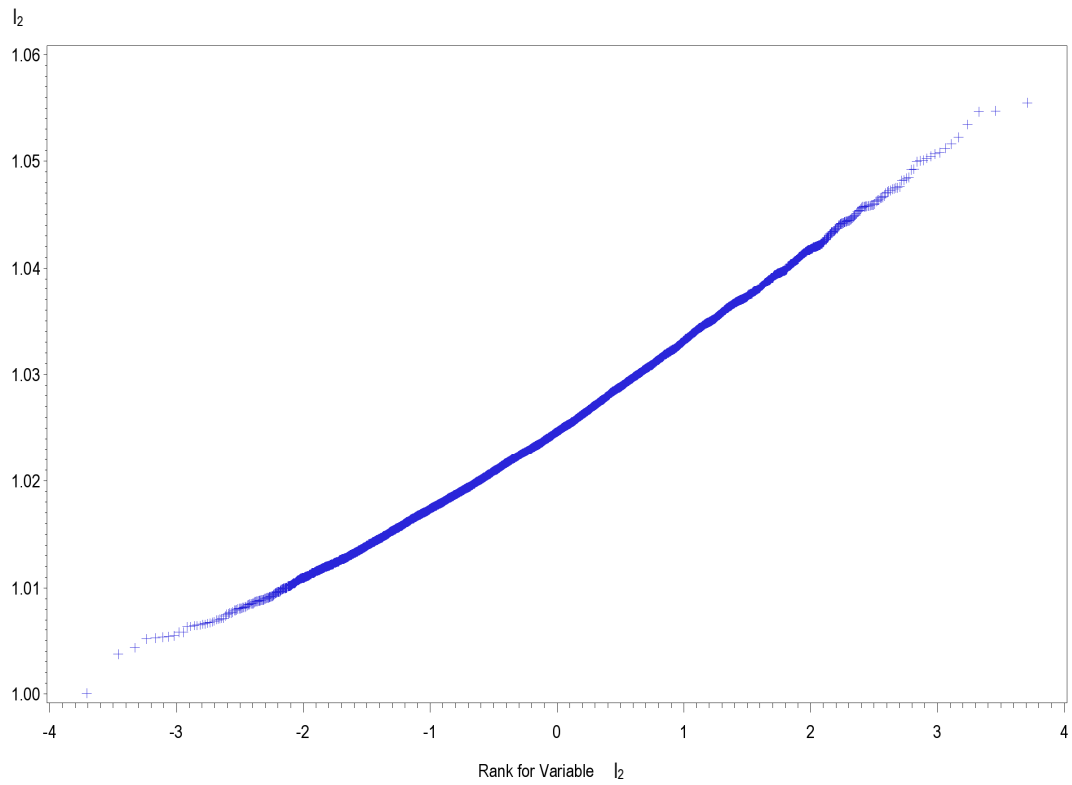
nsub=600 nvar=6



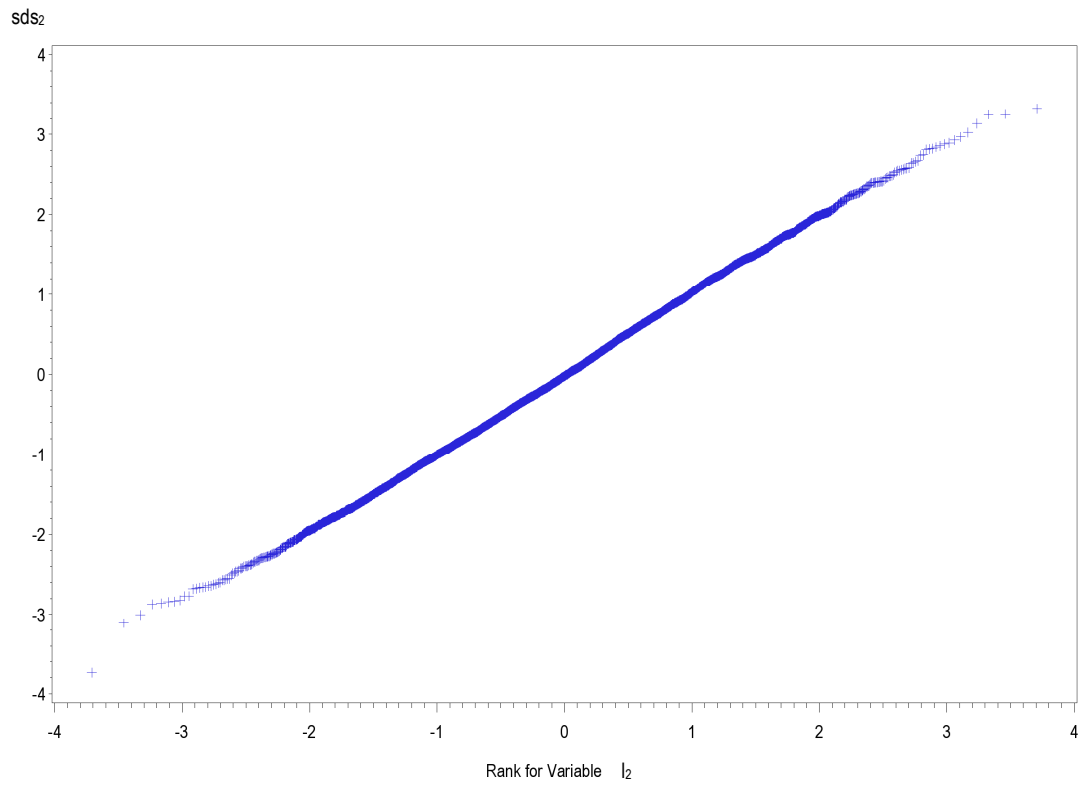
nsub=600 nvar=6



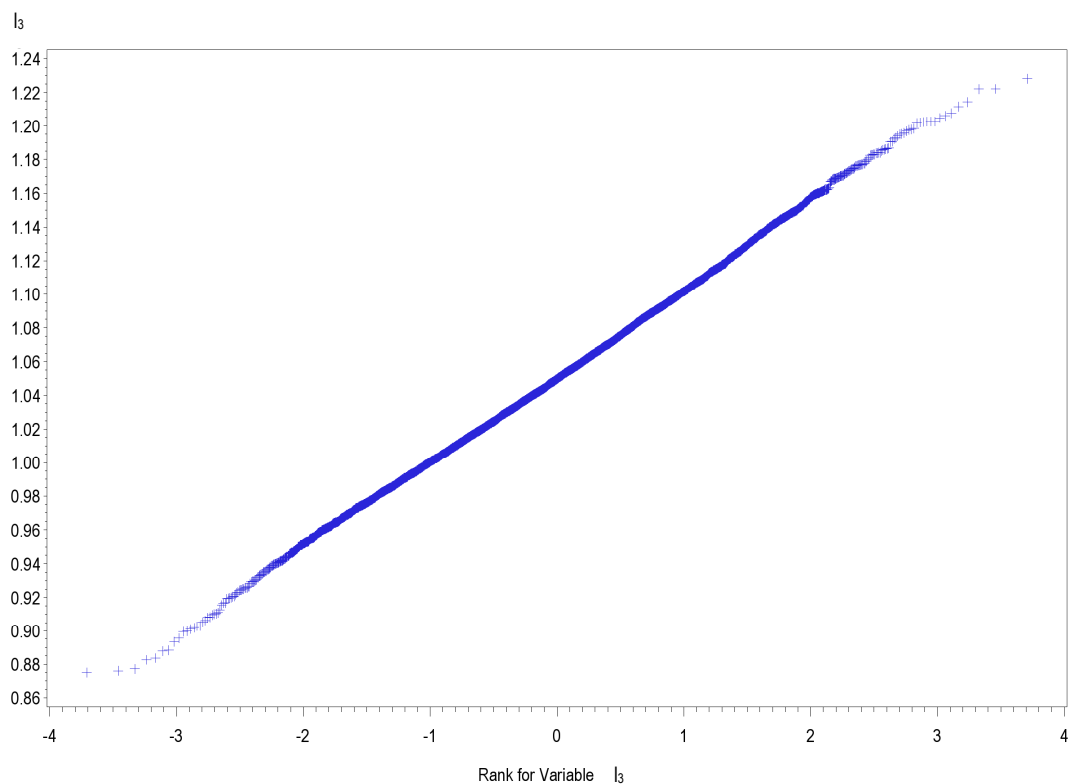
nsub=4800 nvar=6



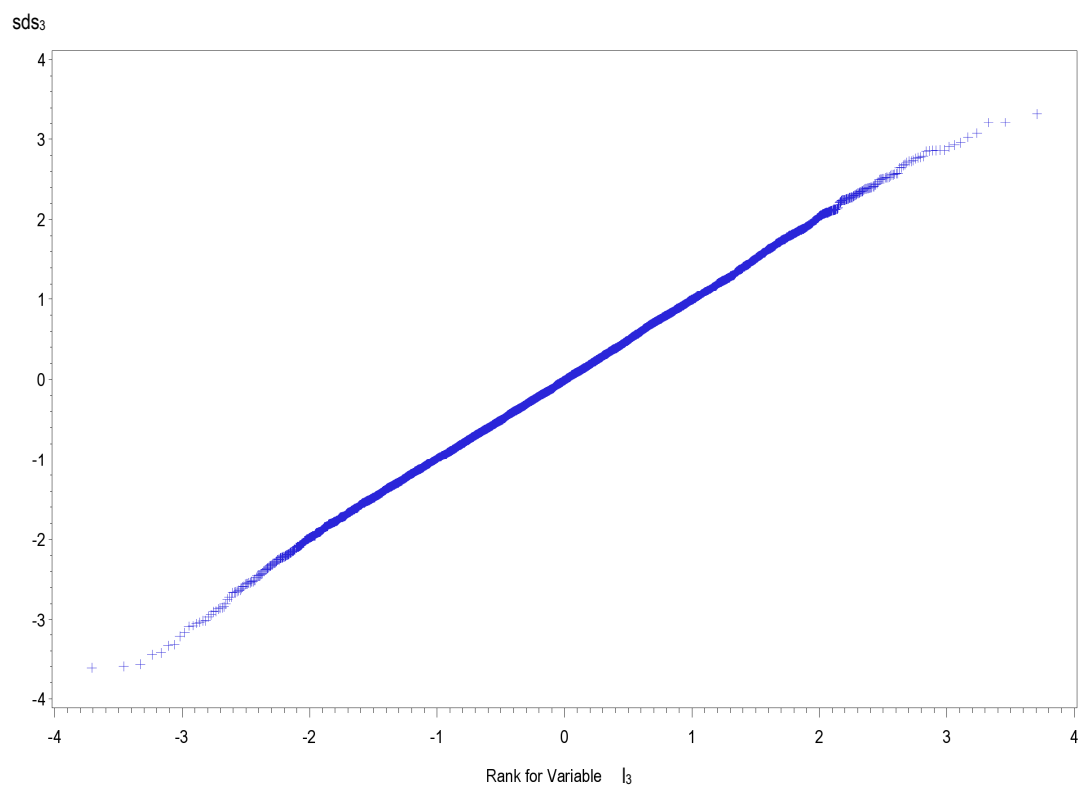
nsub=4800 nvar=6



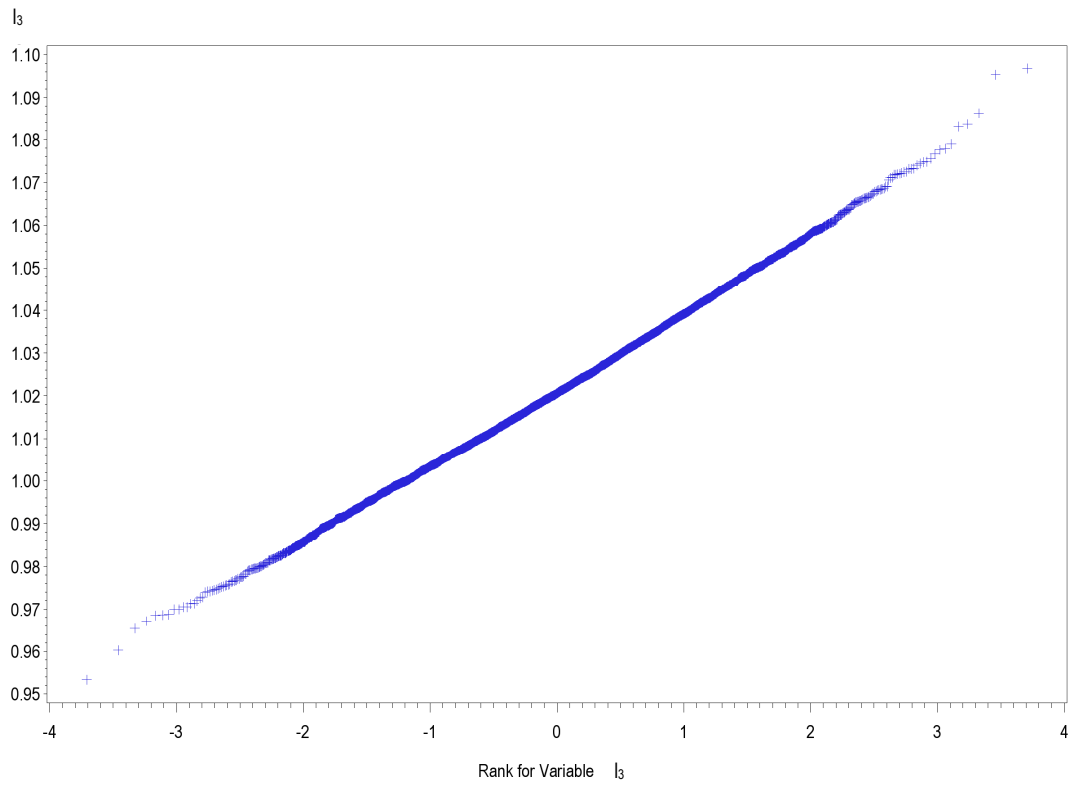
nsub=75 nvar=6



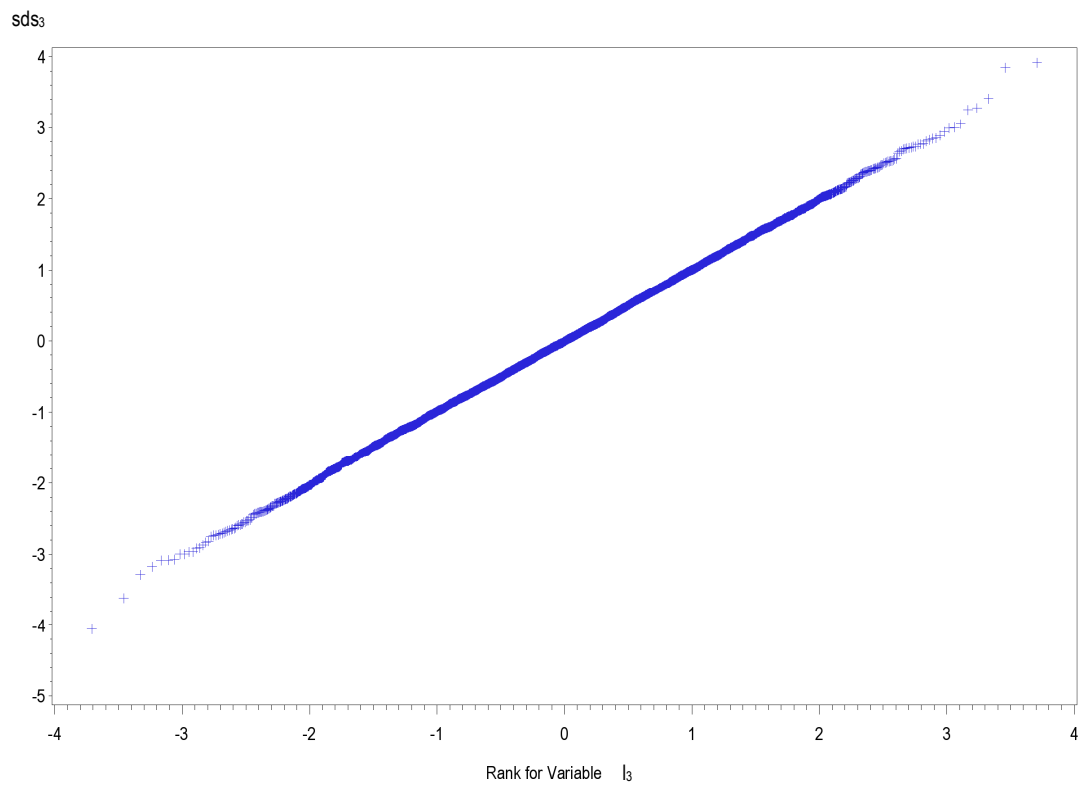
nsub=75 nvar=6



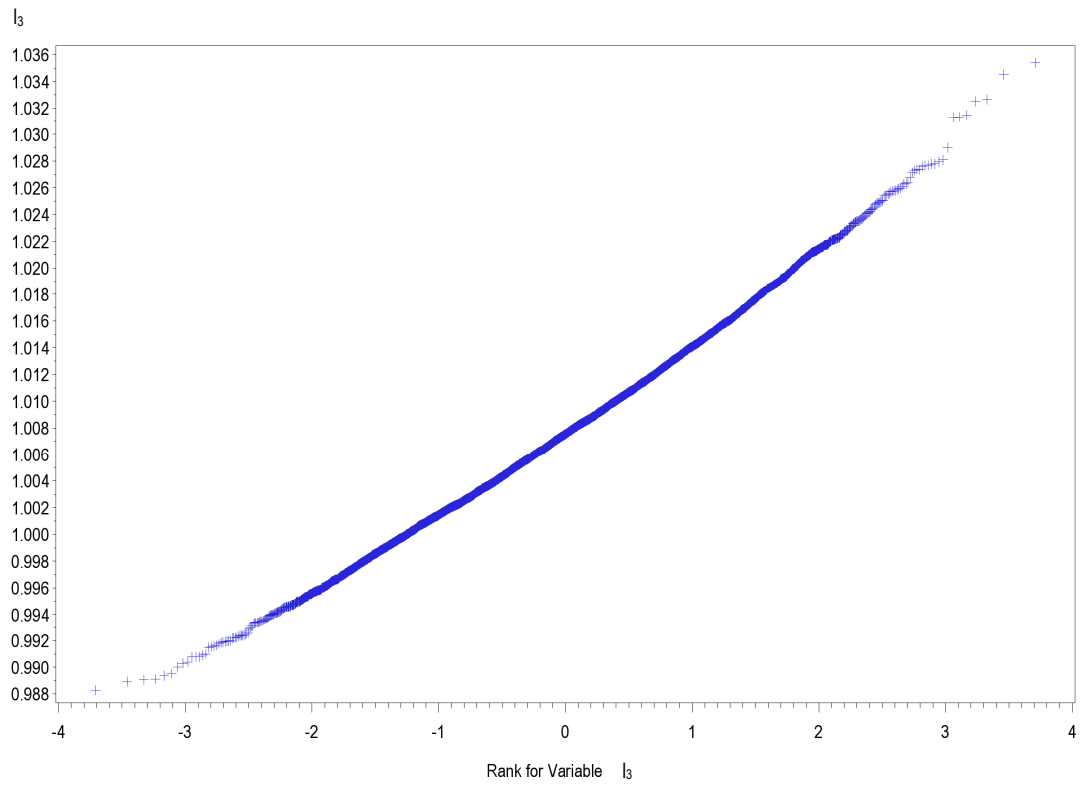
nsub=600 nvar=6



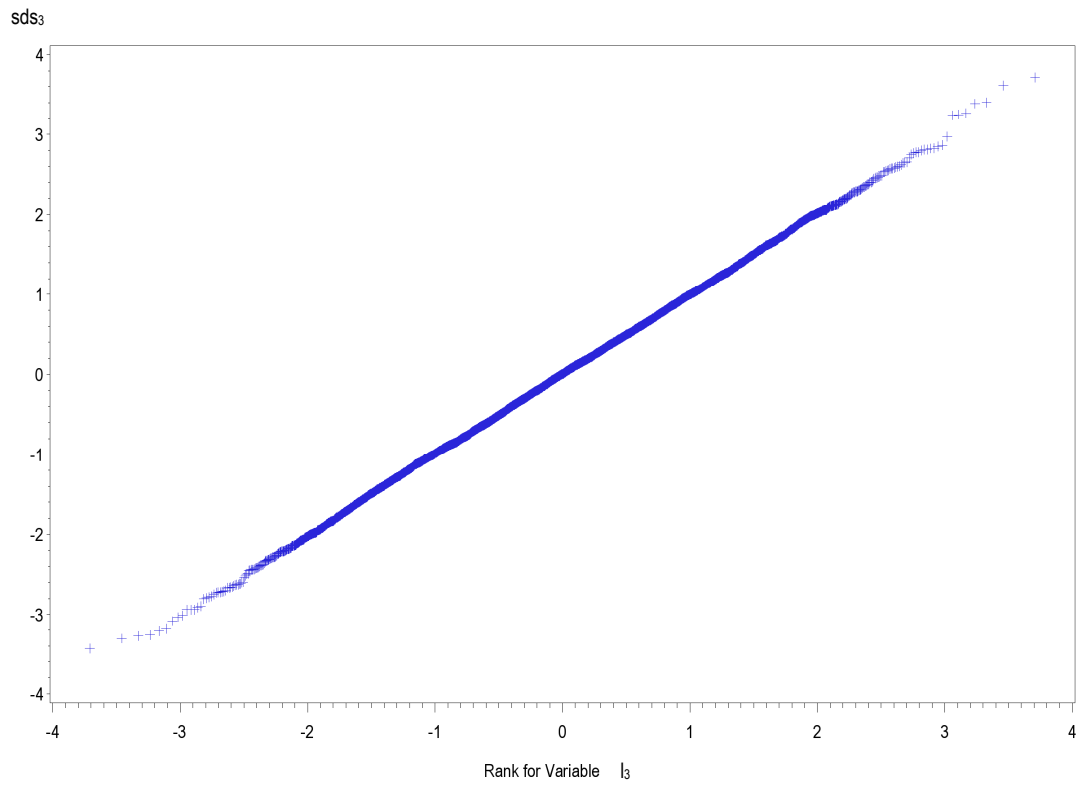
nsub=600 nvar=6



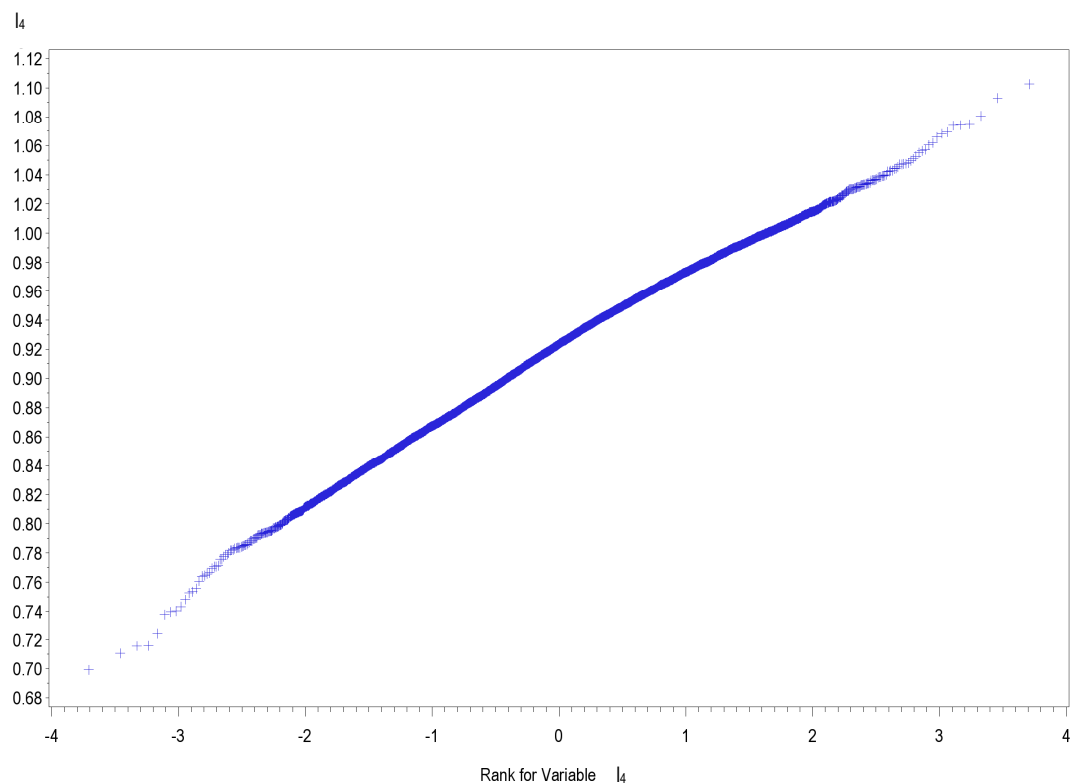
nsub=4800 nvar=6



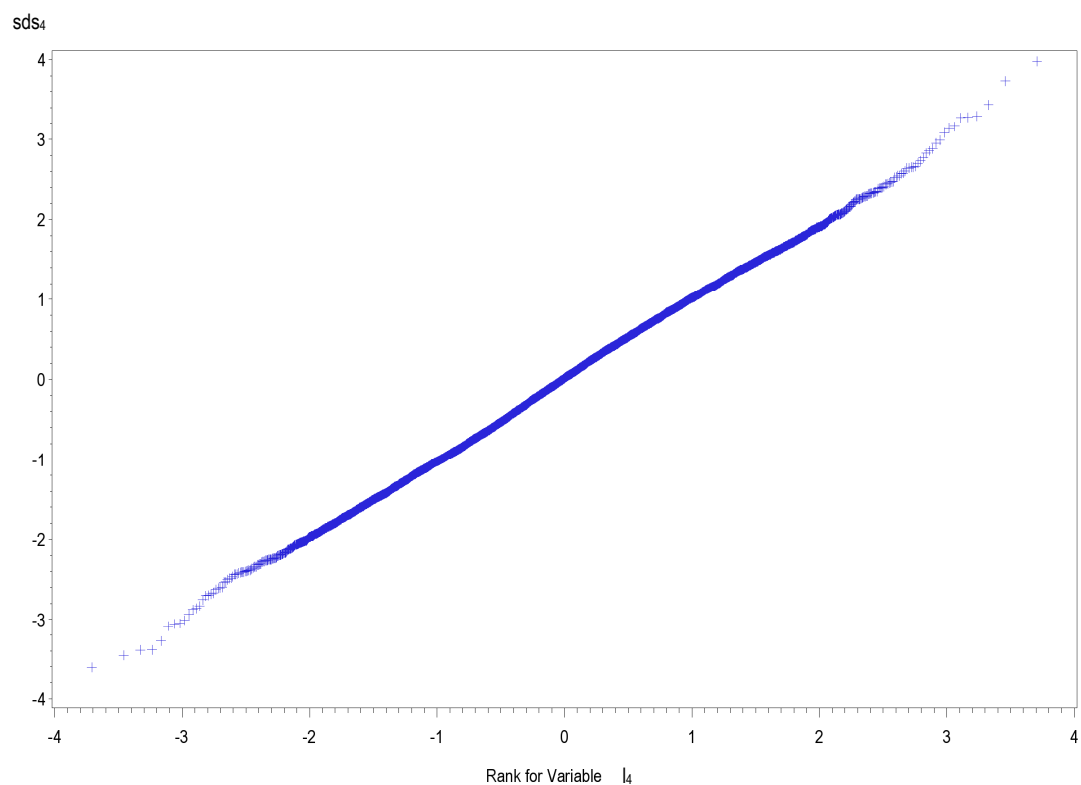
nsub=4800 nvar=6



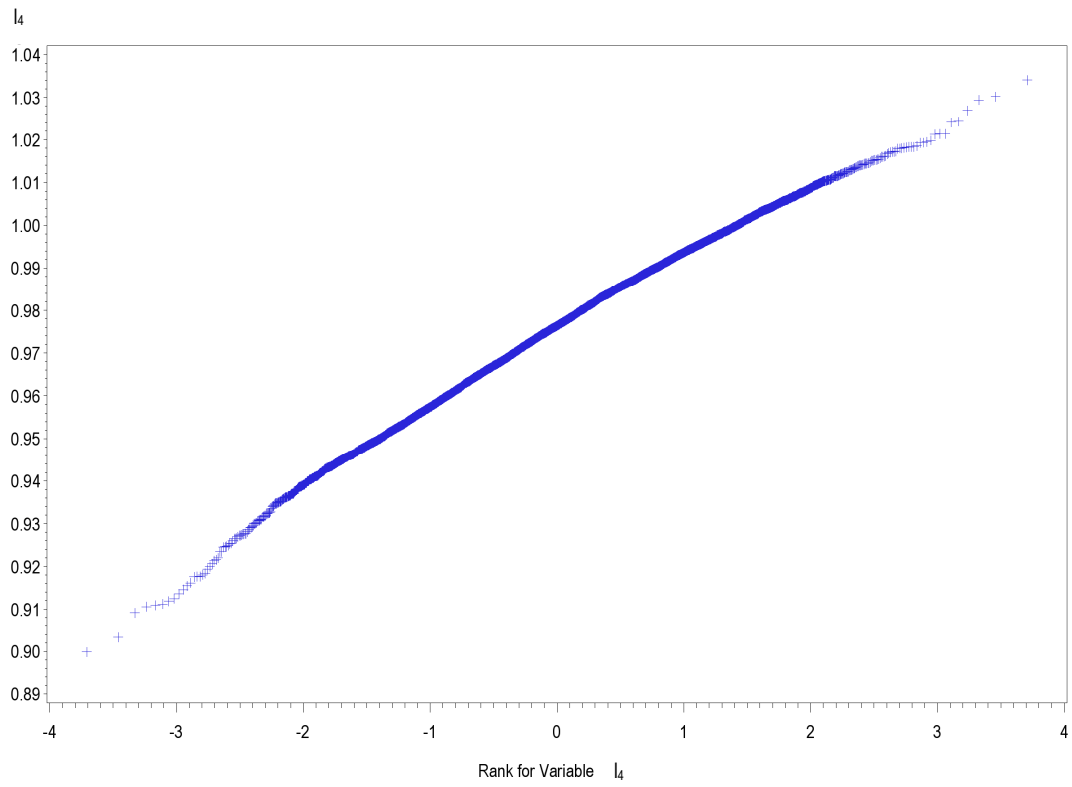
nsub=75 nvar=6



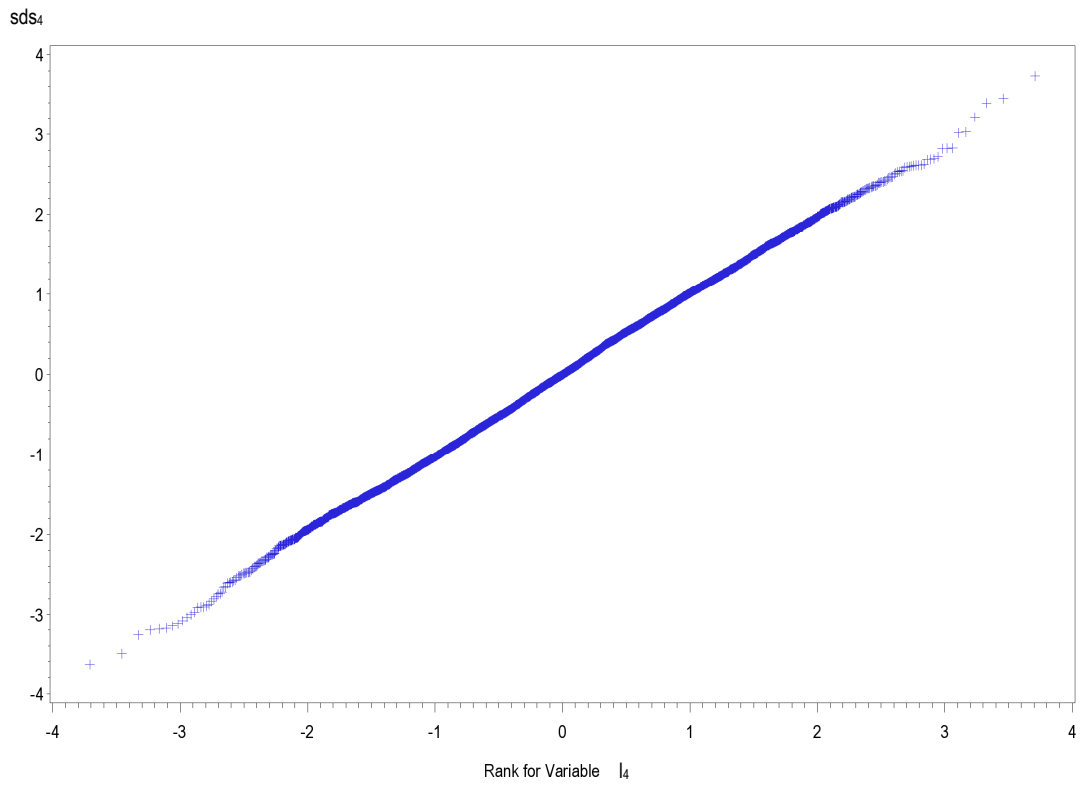
nsub=75 nvar=6



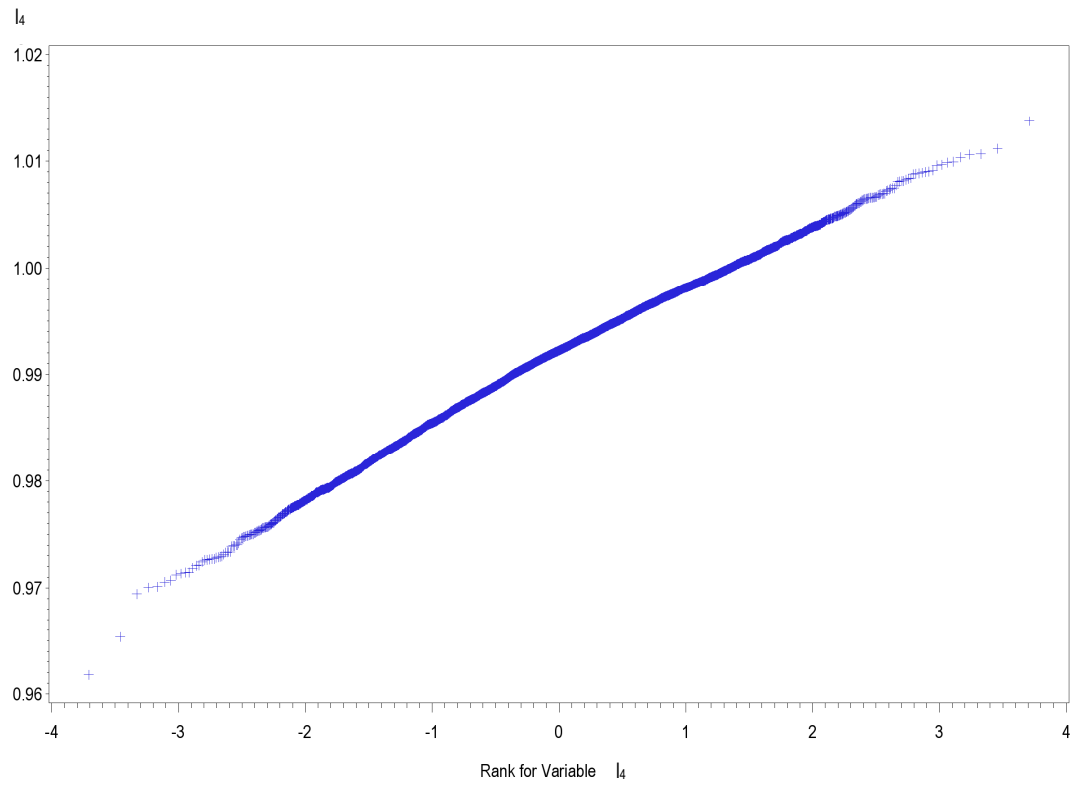
nsub=600 nvar=6



nsub=600 nvar=6



nsub=4800 nvar=6



nsub=4800 nvar=6

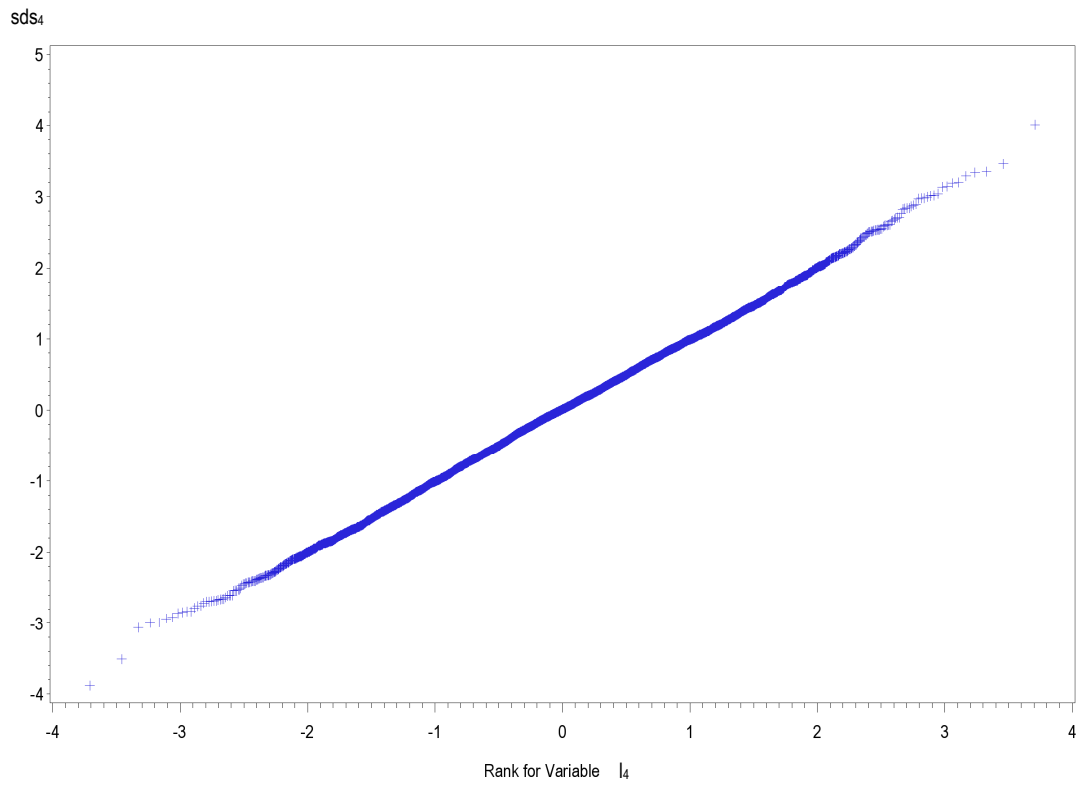
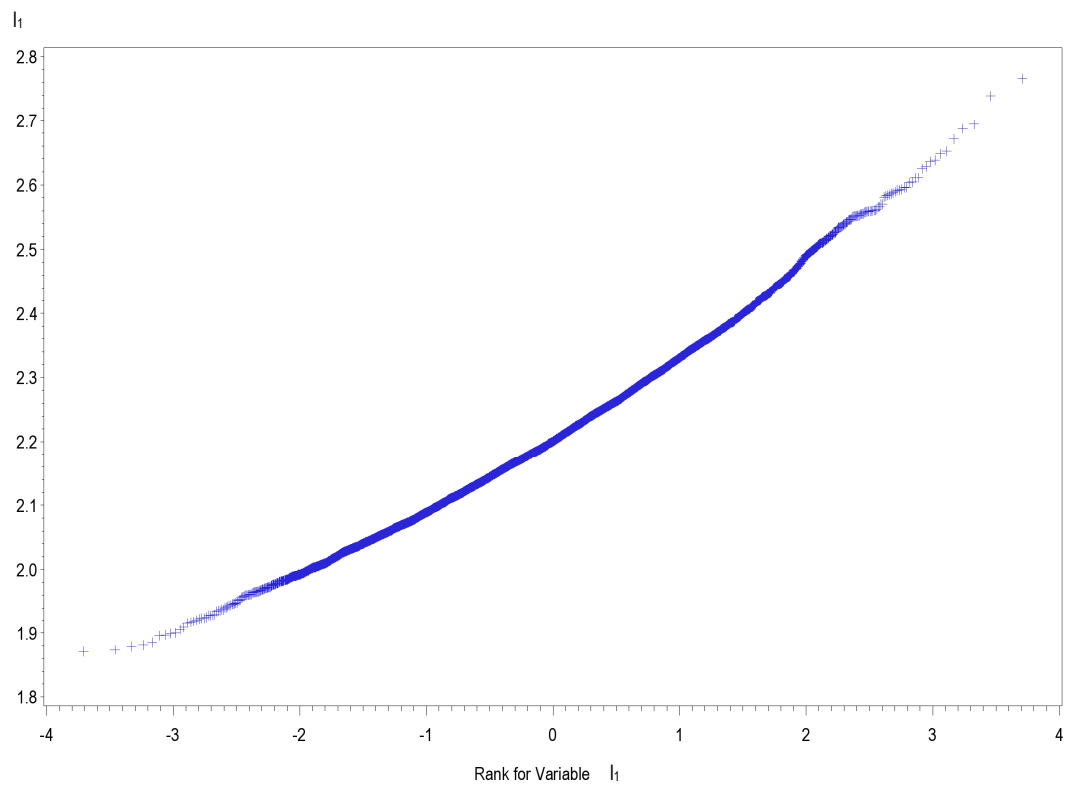
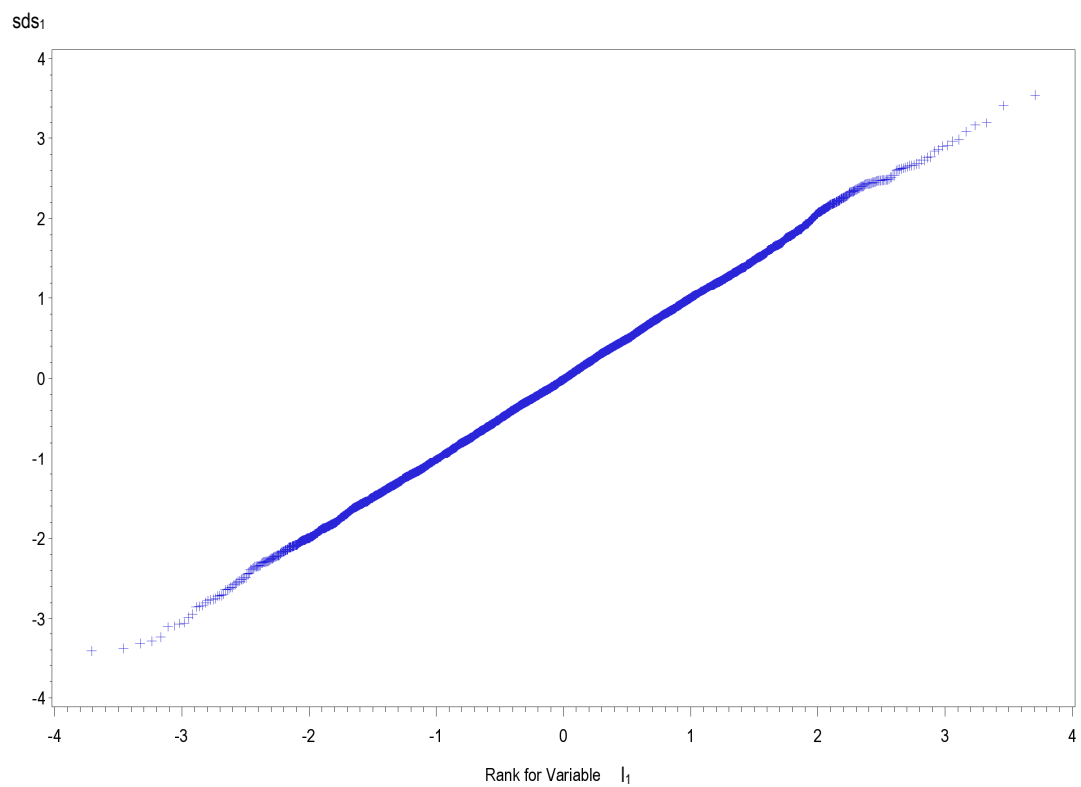


Figura A25-A48. Q-Qplot per l_1-l_4 e sds_1-sds_4 relativi agli scenari con 24 variabili per 75, 600 e 4800 soggetti

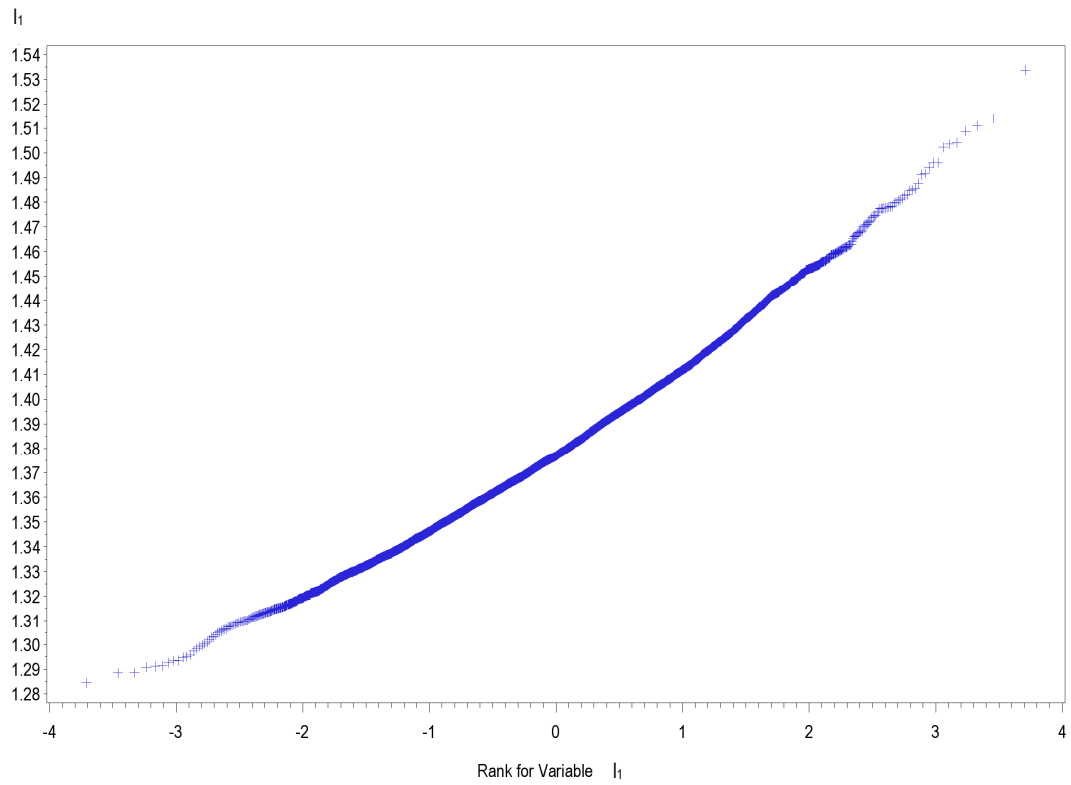
nsub=75 nvar=24



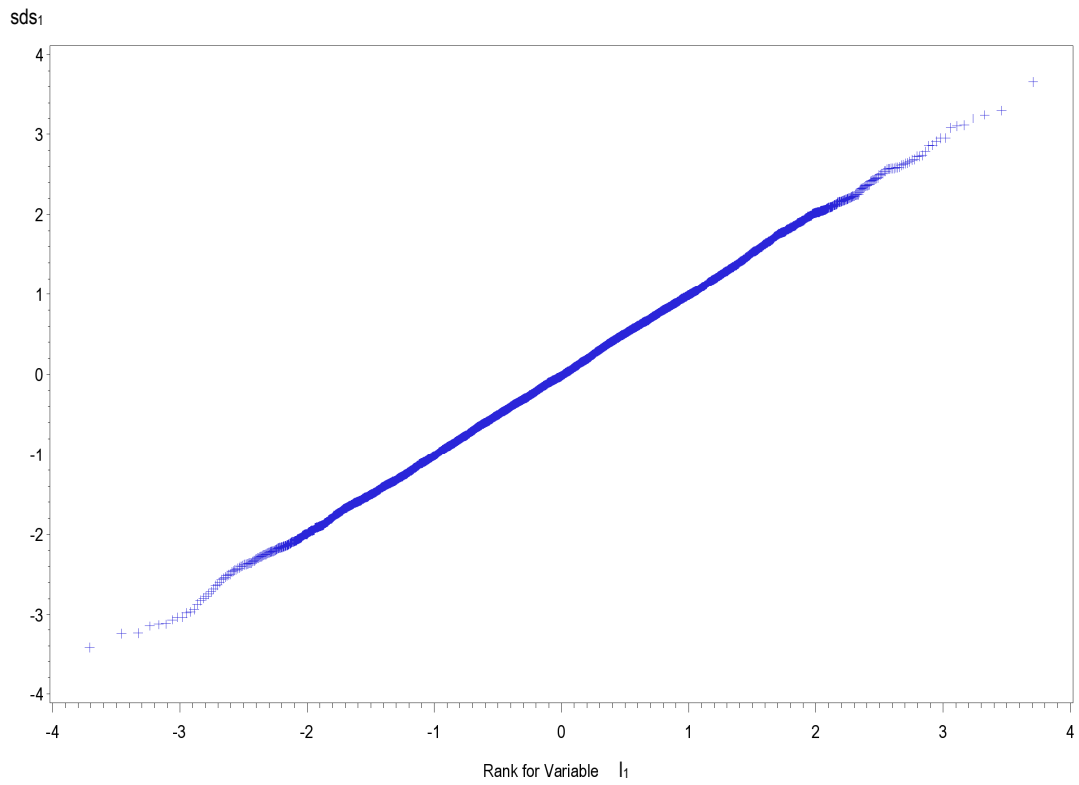
nsub=75 nvar=24



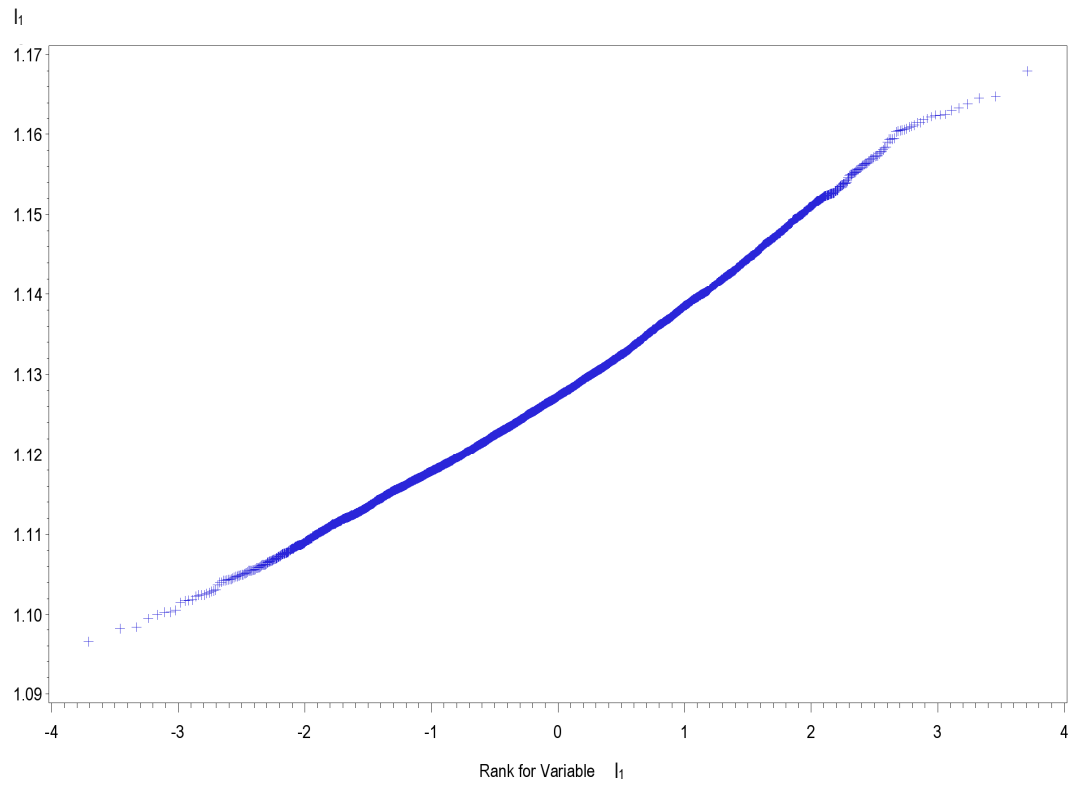
nsub=600 nvar=24



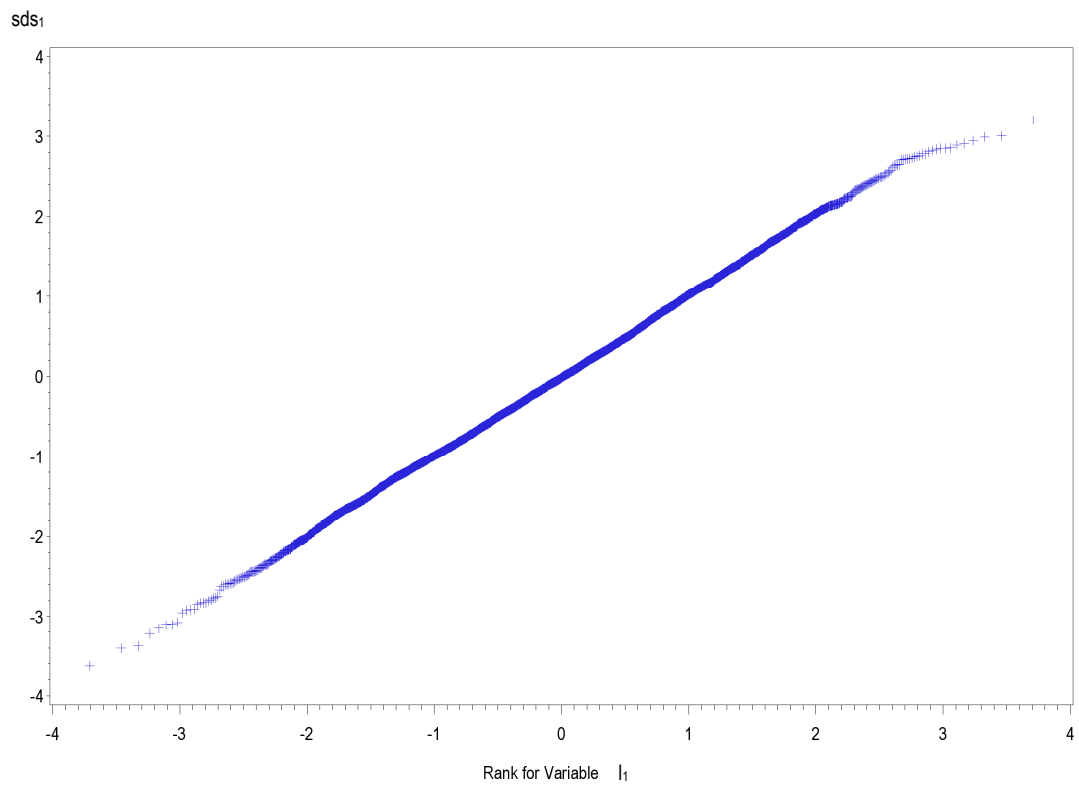
nsub=600 nvar=24



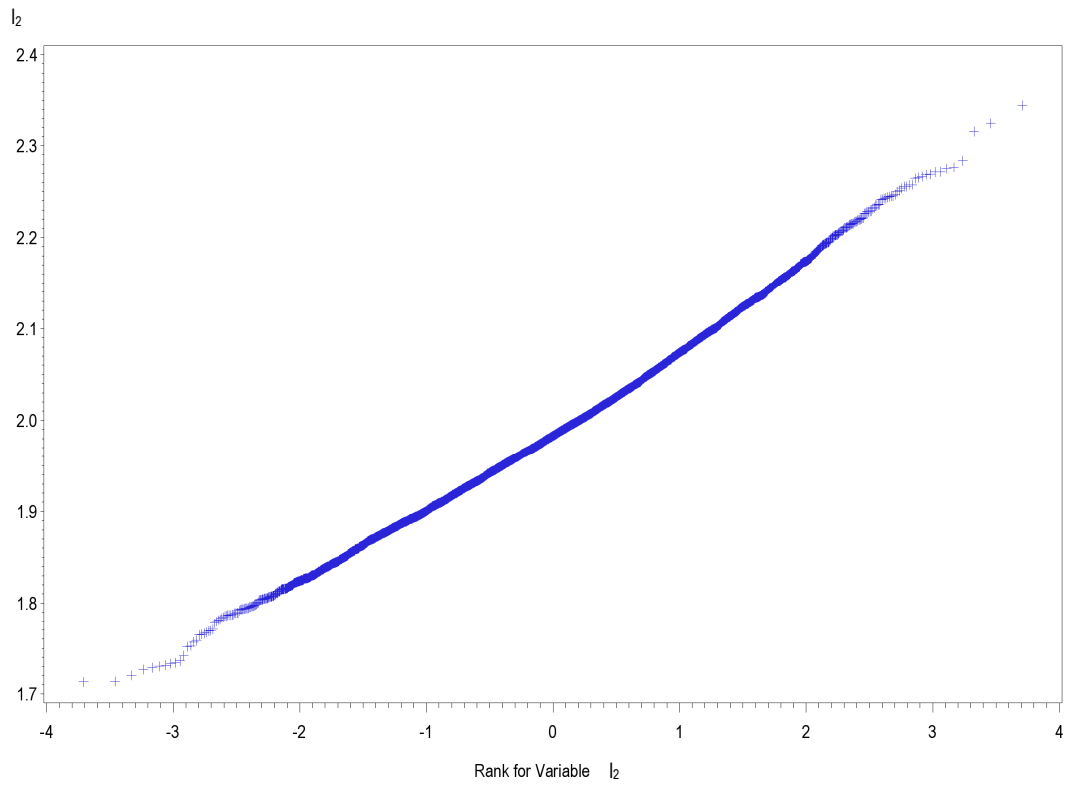
nsub=4800 nvar=24



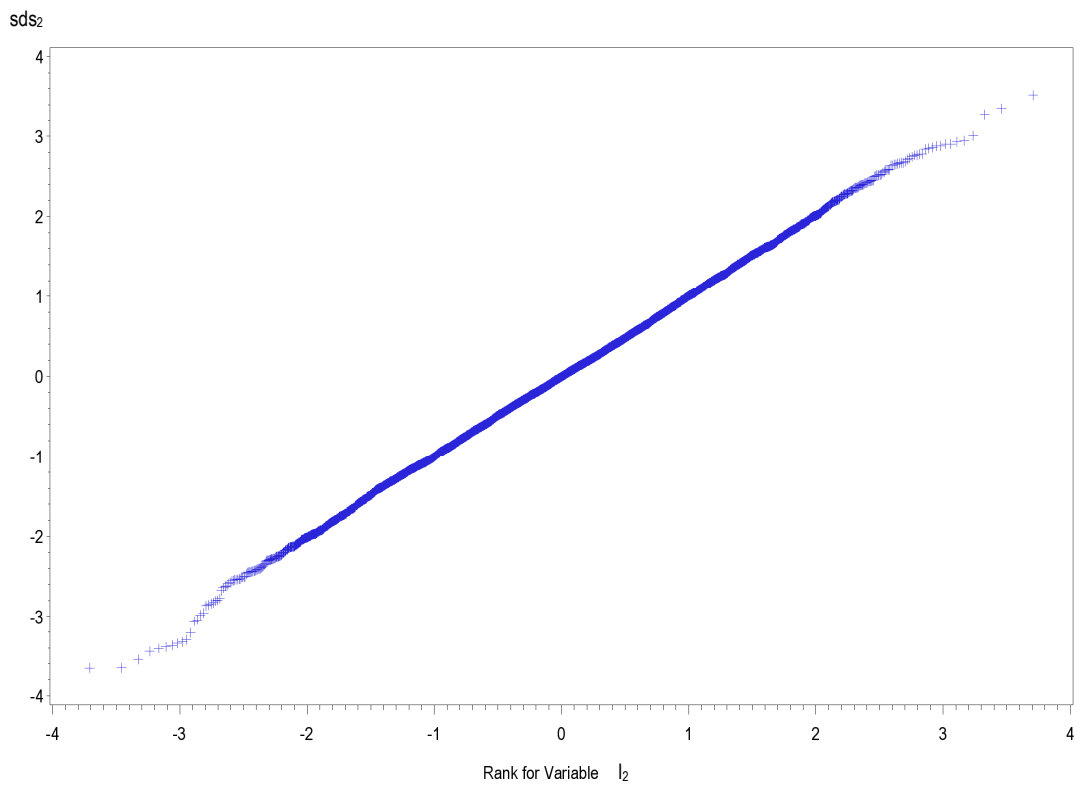
nsub=4800 nvar=24



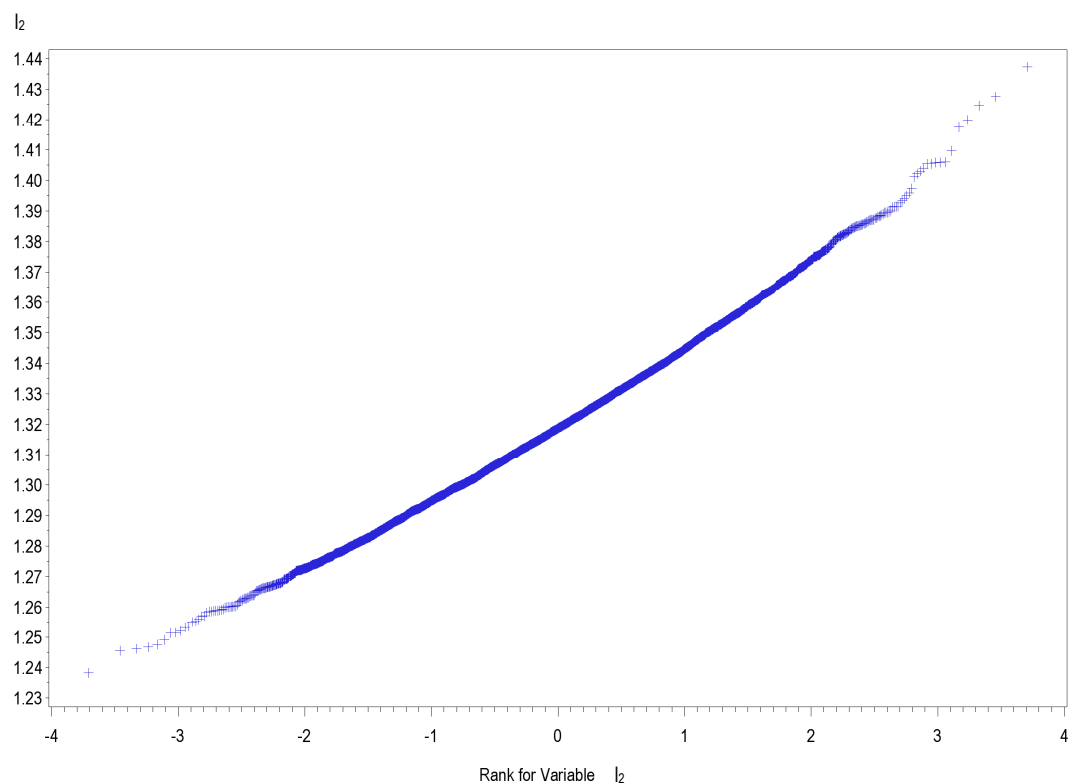
nsub=75 nvar=24



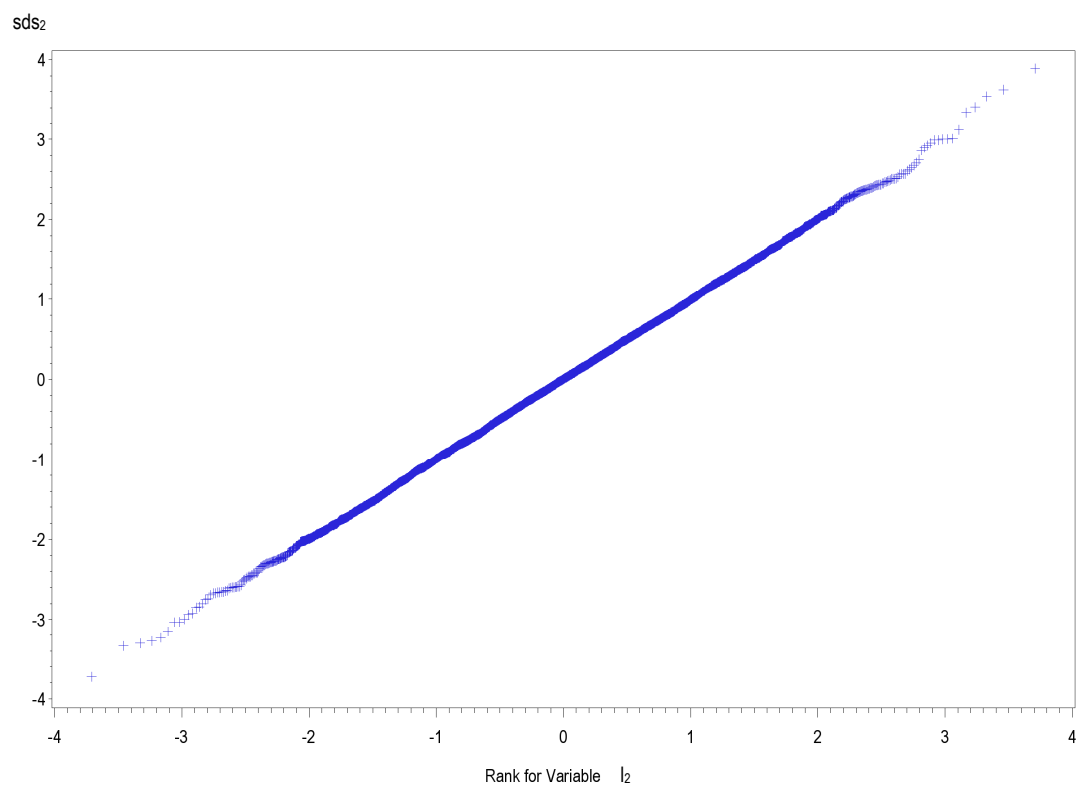
nsub=75 nvar=24



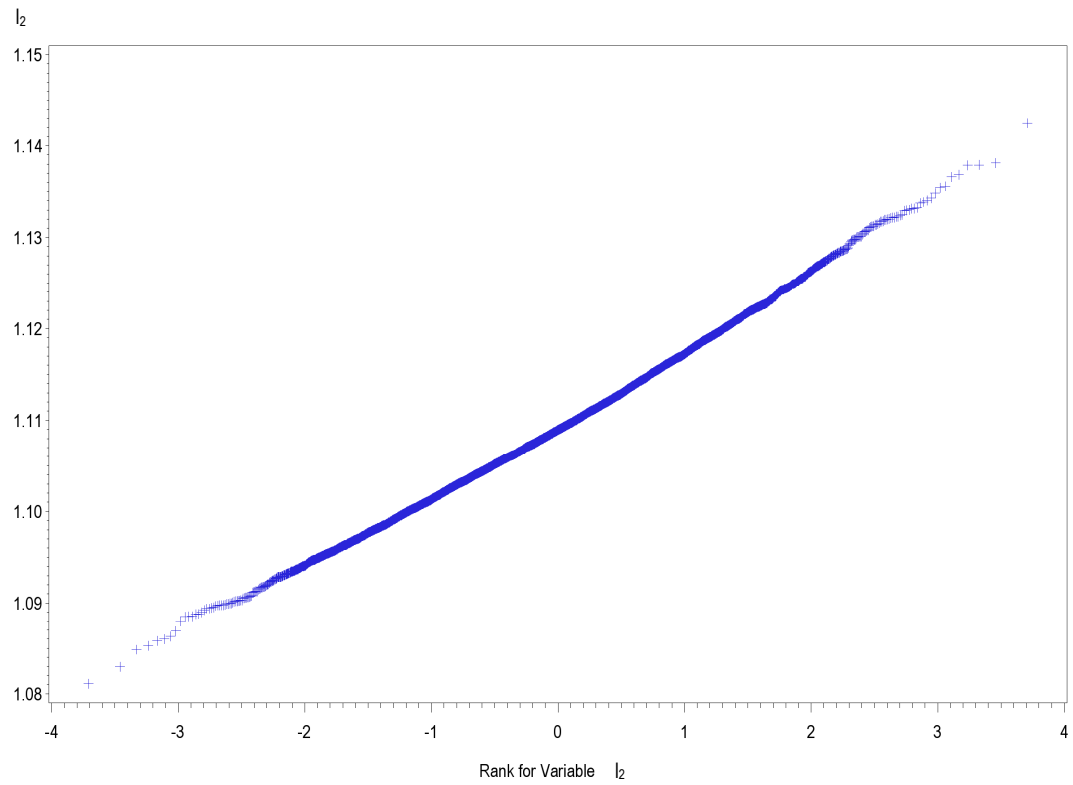
nsub=600 nvar=24



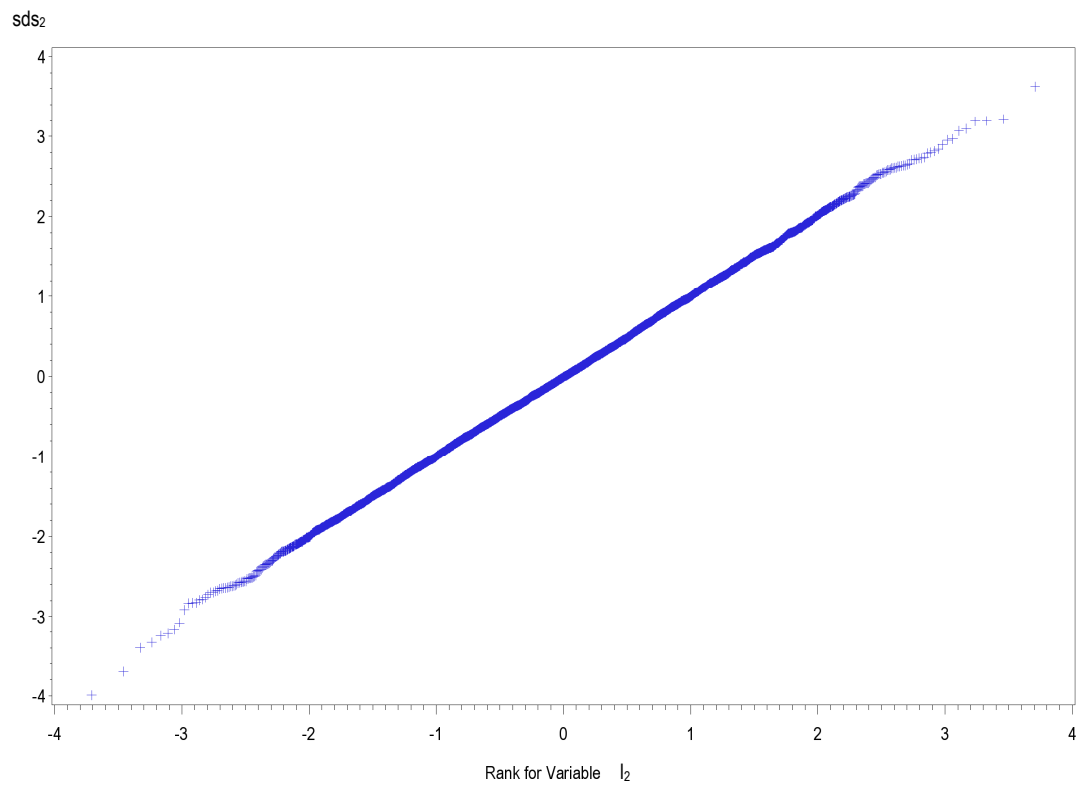
nsub=600 nvar=24



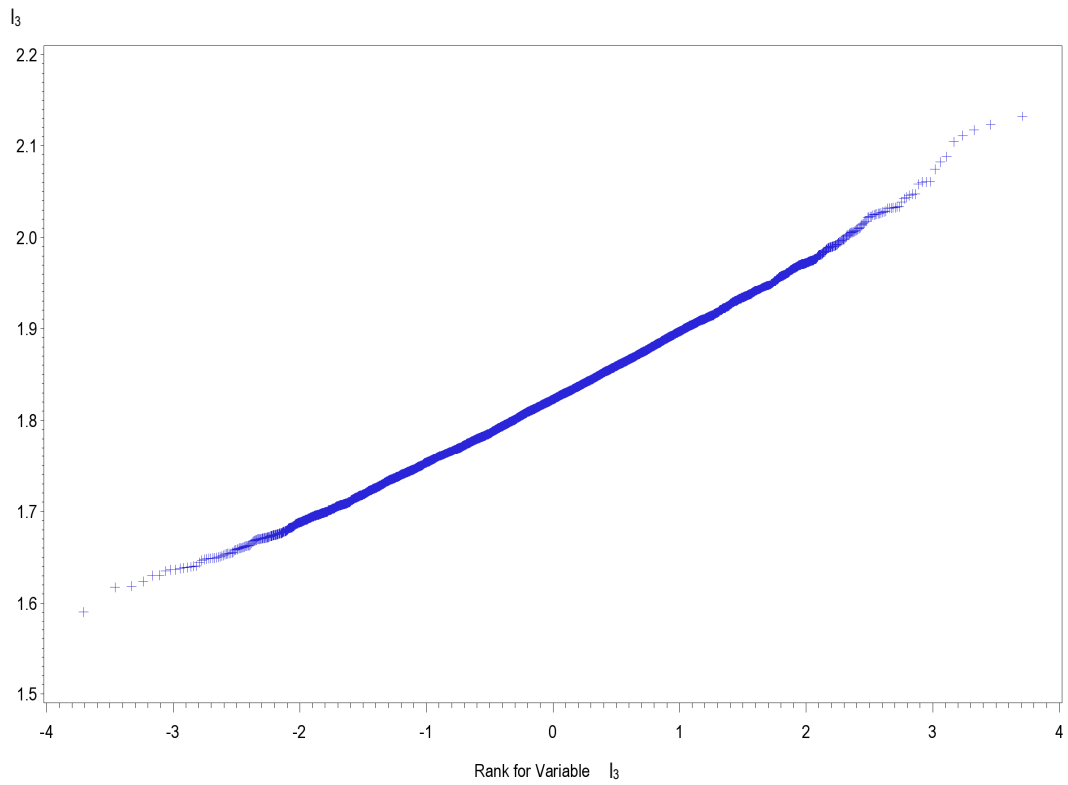
nsub=4800 nvar=24



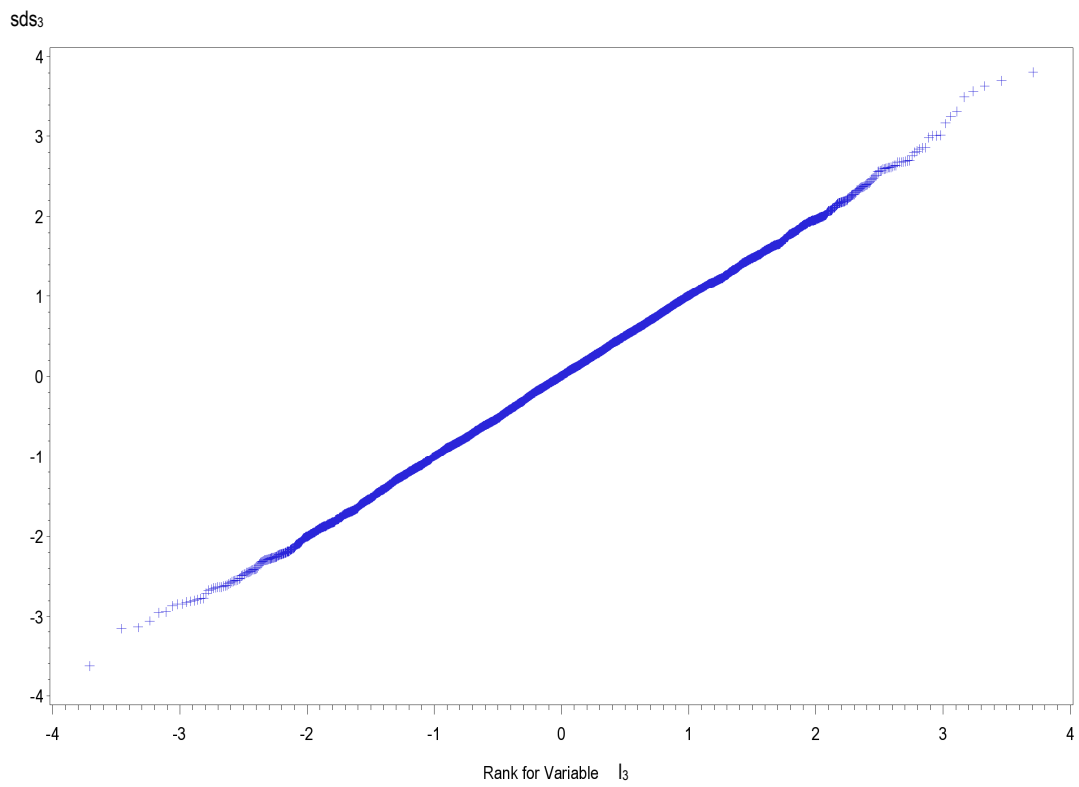
nsub=4800 nvar=24



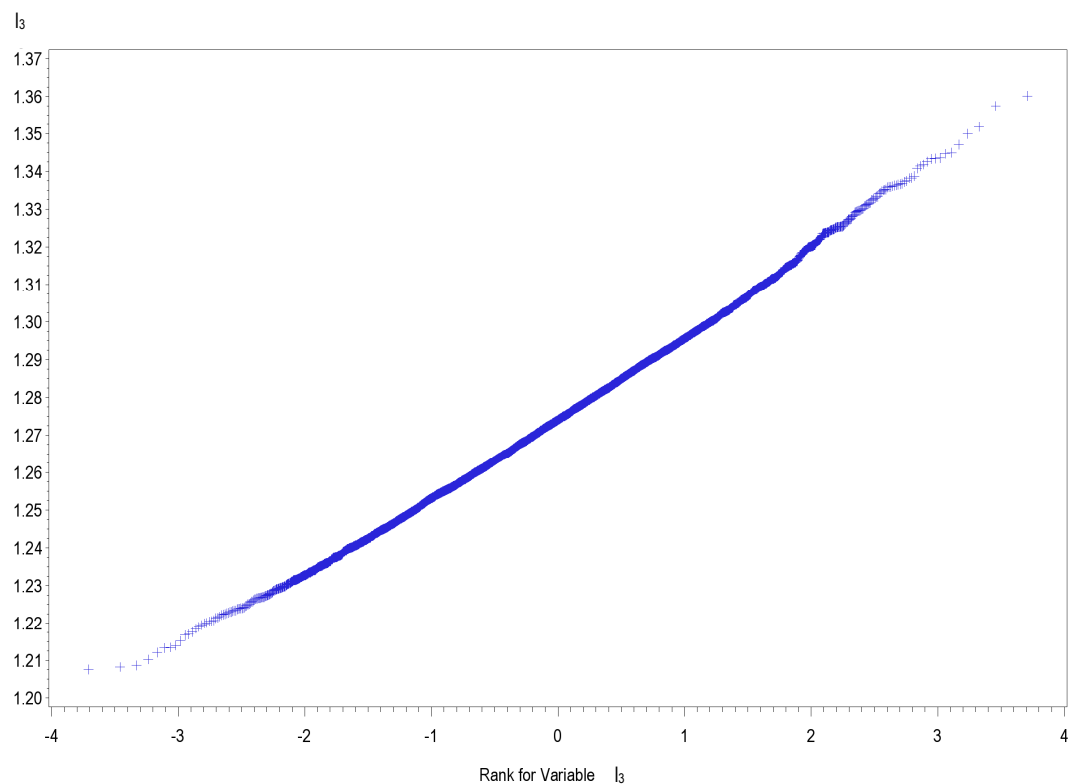
nsub=75 nvar=24



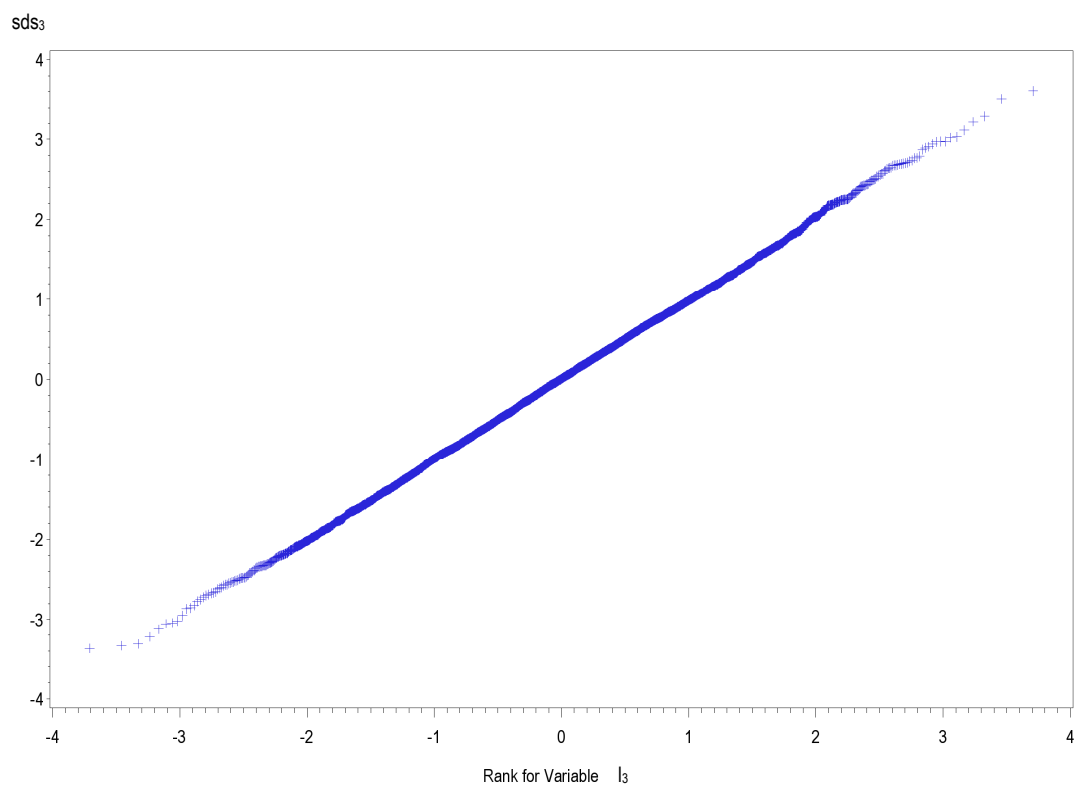
nsub=75 nvar=24



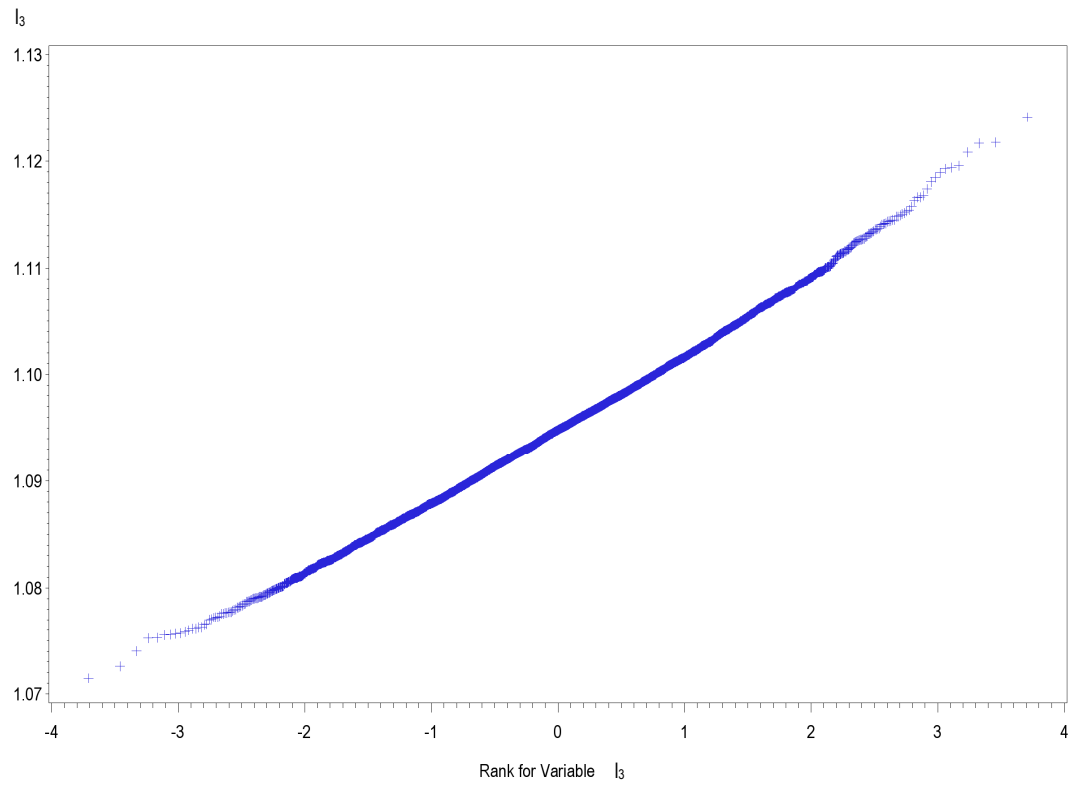
nsub=600 nvar=24



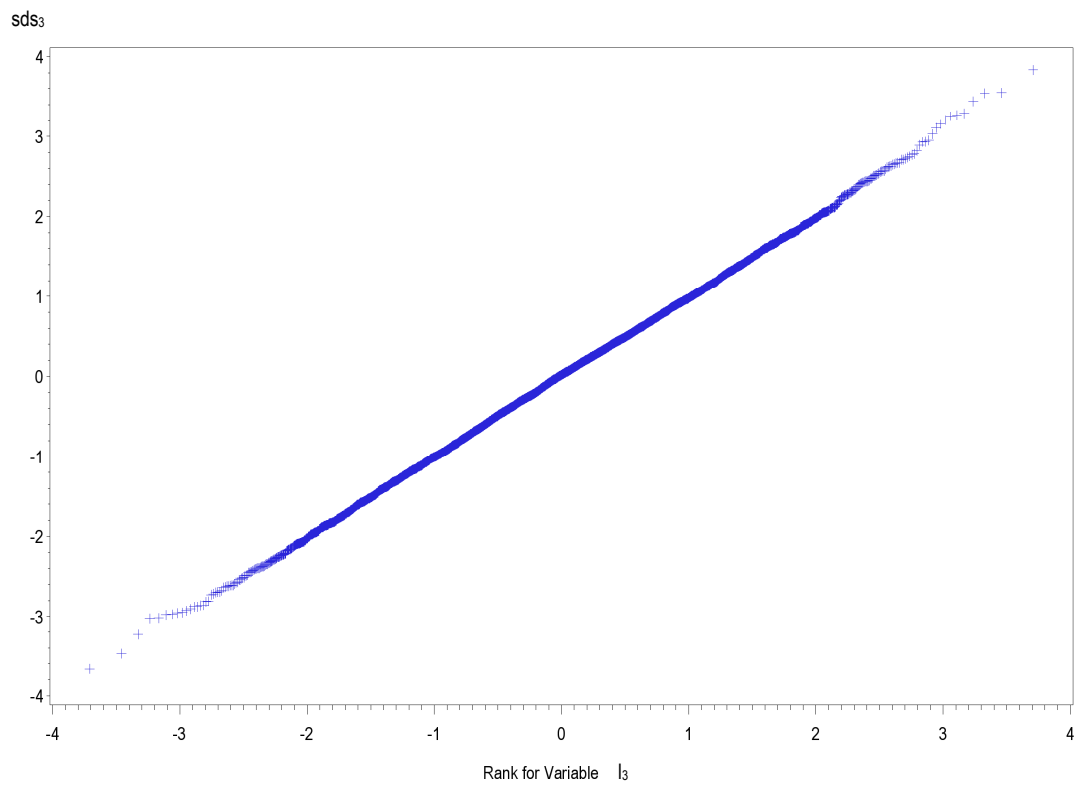
nsub=600 nvar=24



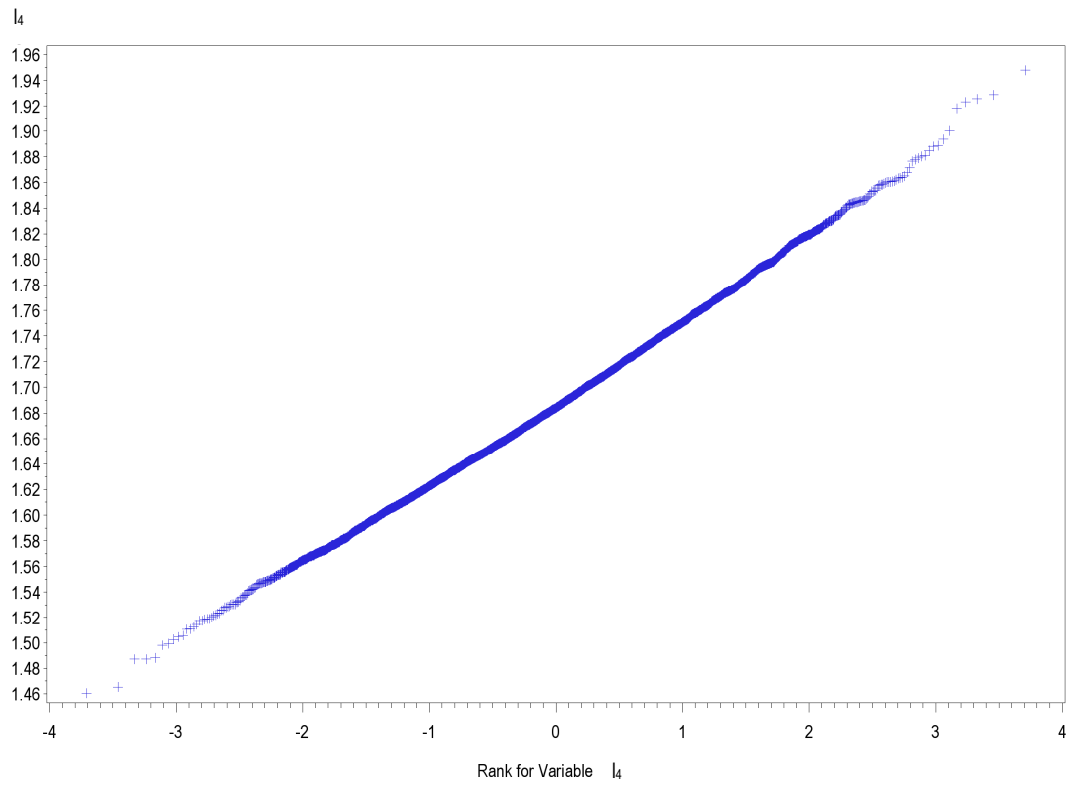
nsub=4800 nvar=24



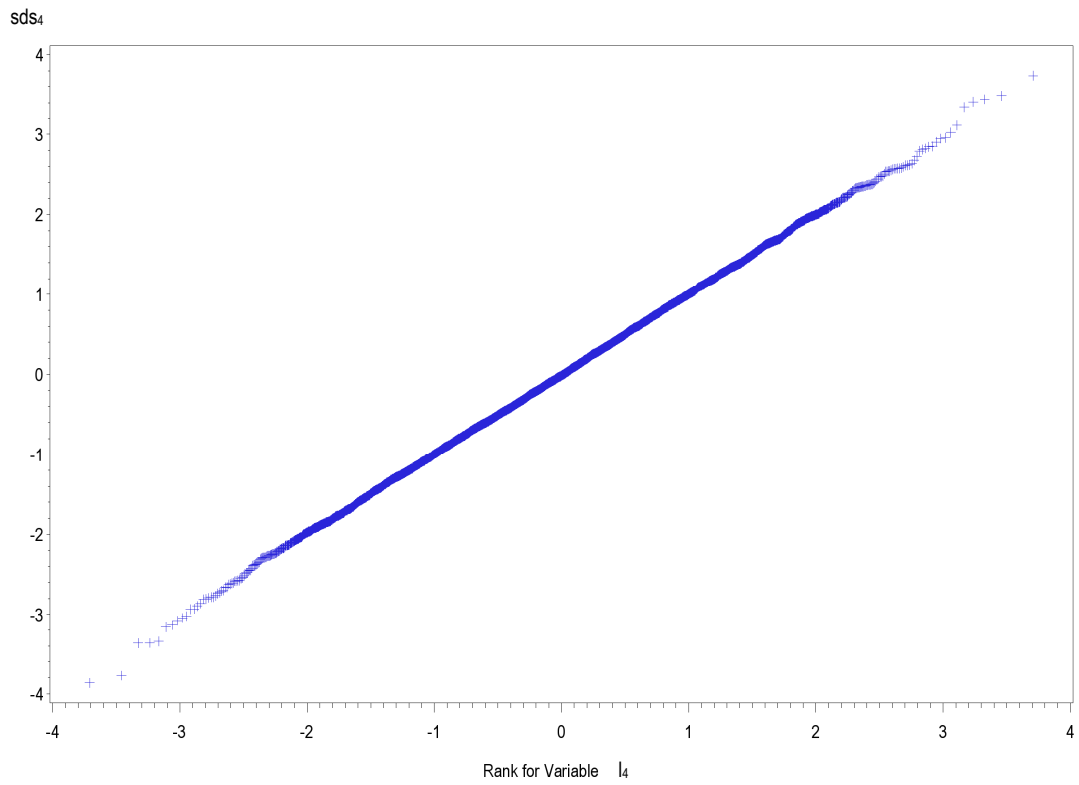
nsub=4800 nvar=24



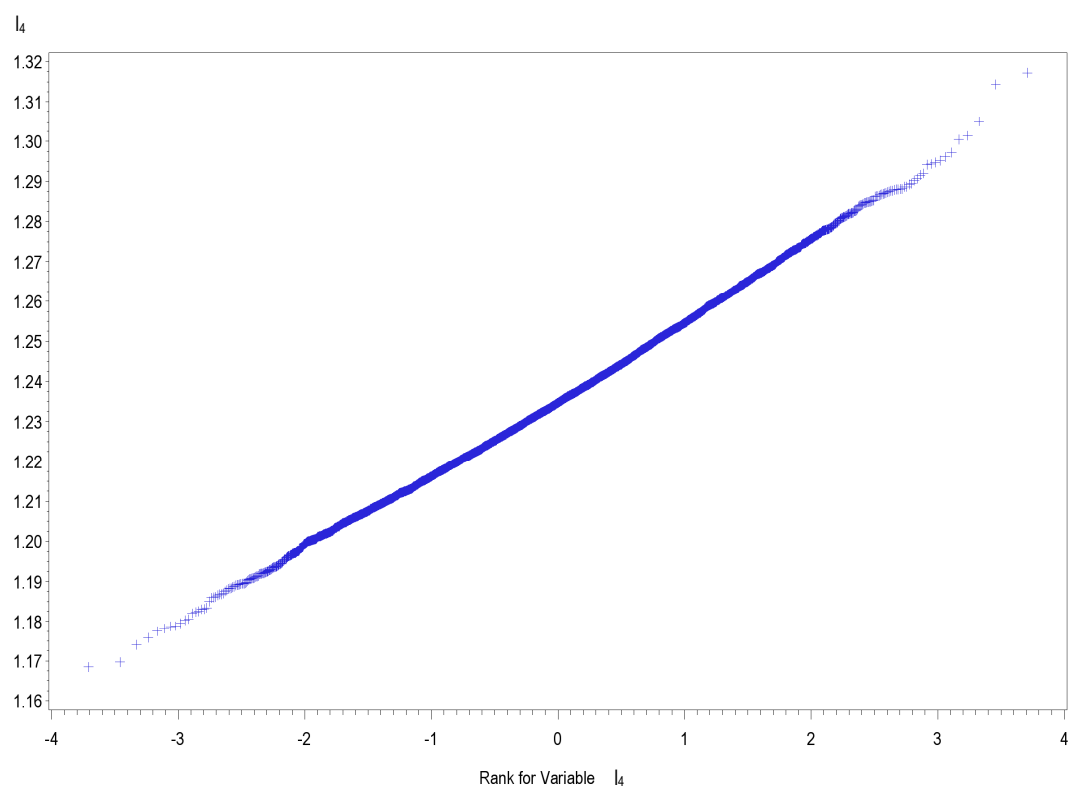
nsub=75 nvar=24



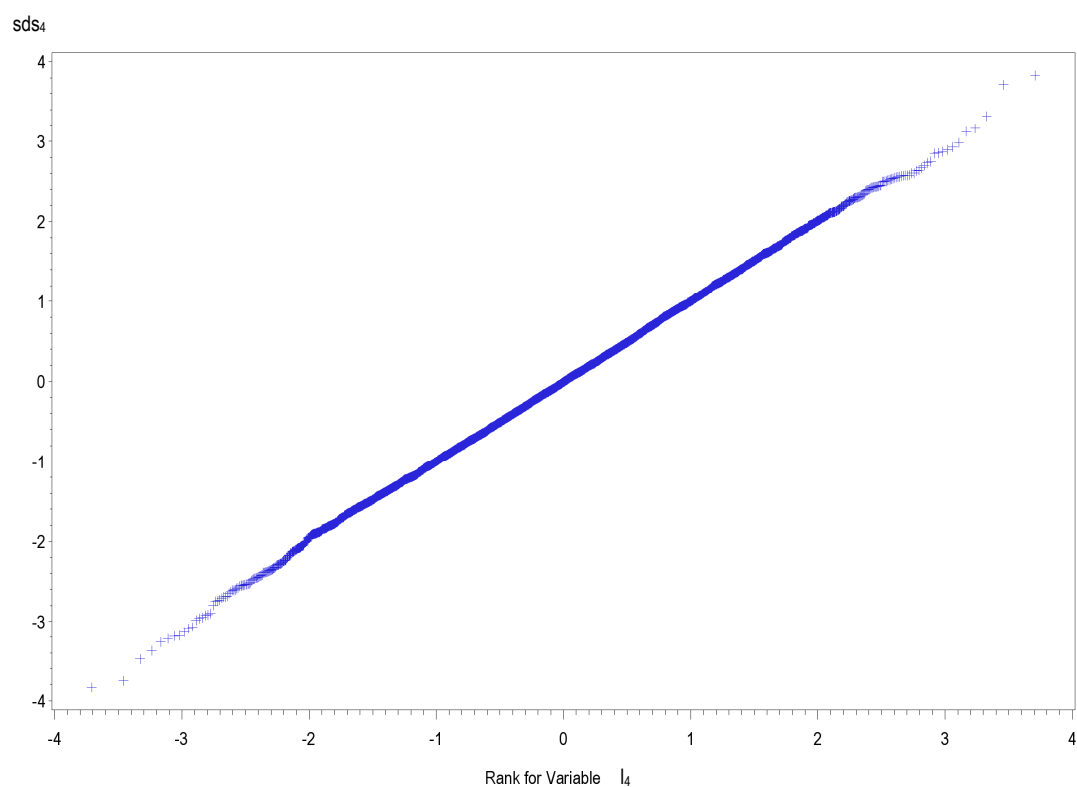
nsub=75 nvar=24



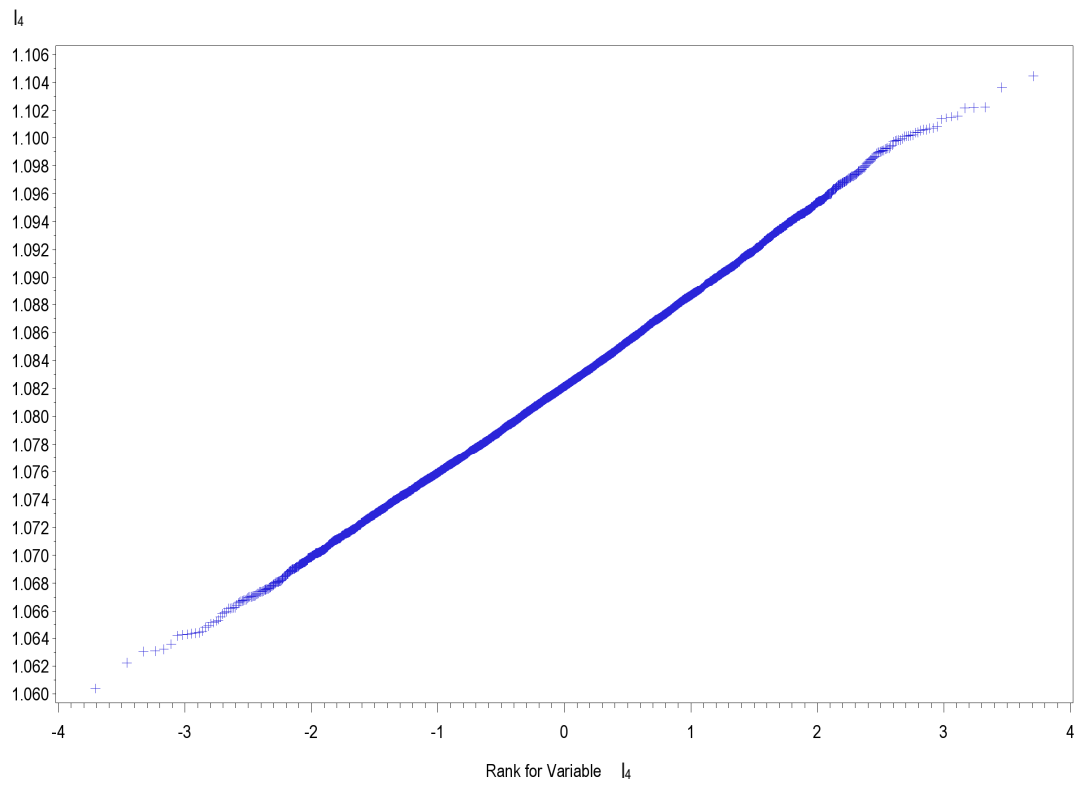
nsub=600 nvar=24



nsub=600 nvar=24



nsub=4800 nvar=24



nsub=4800 nvar=24

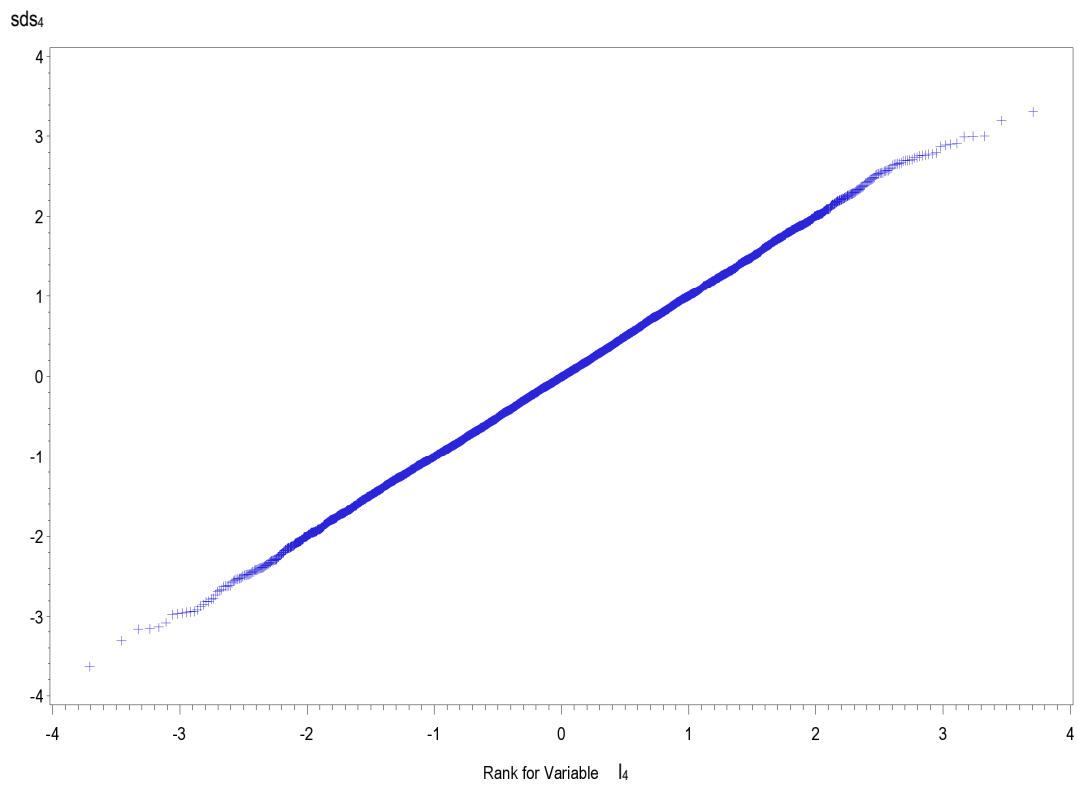
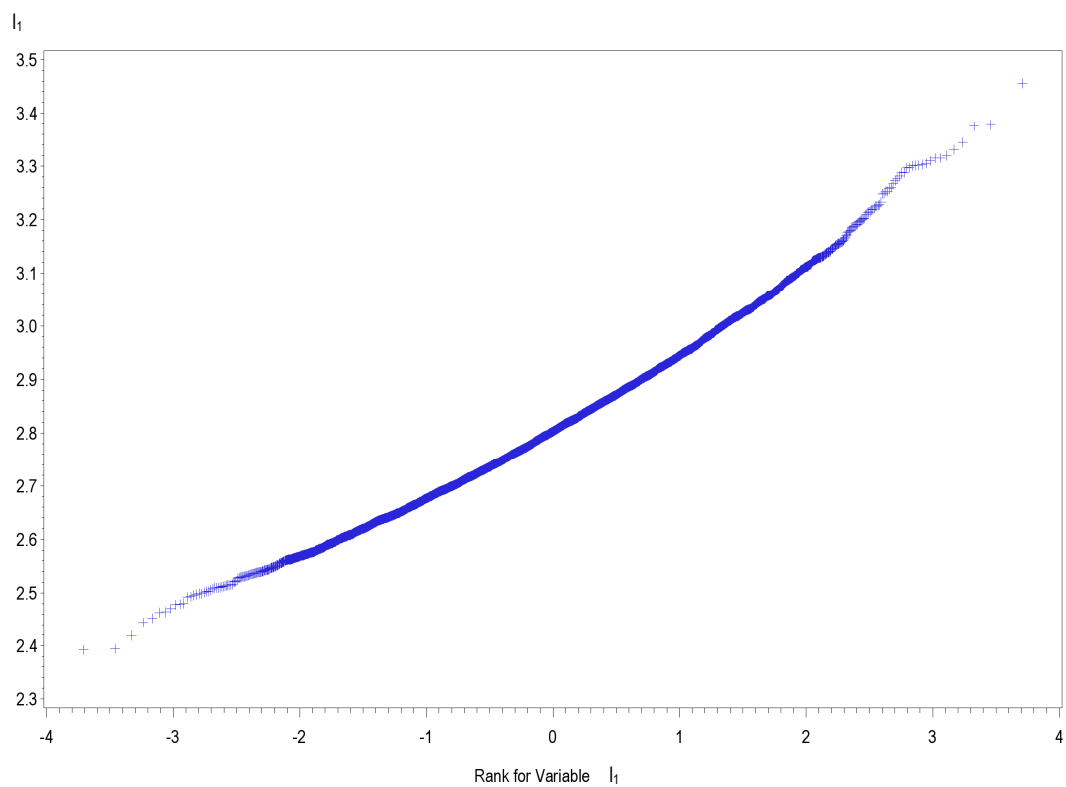
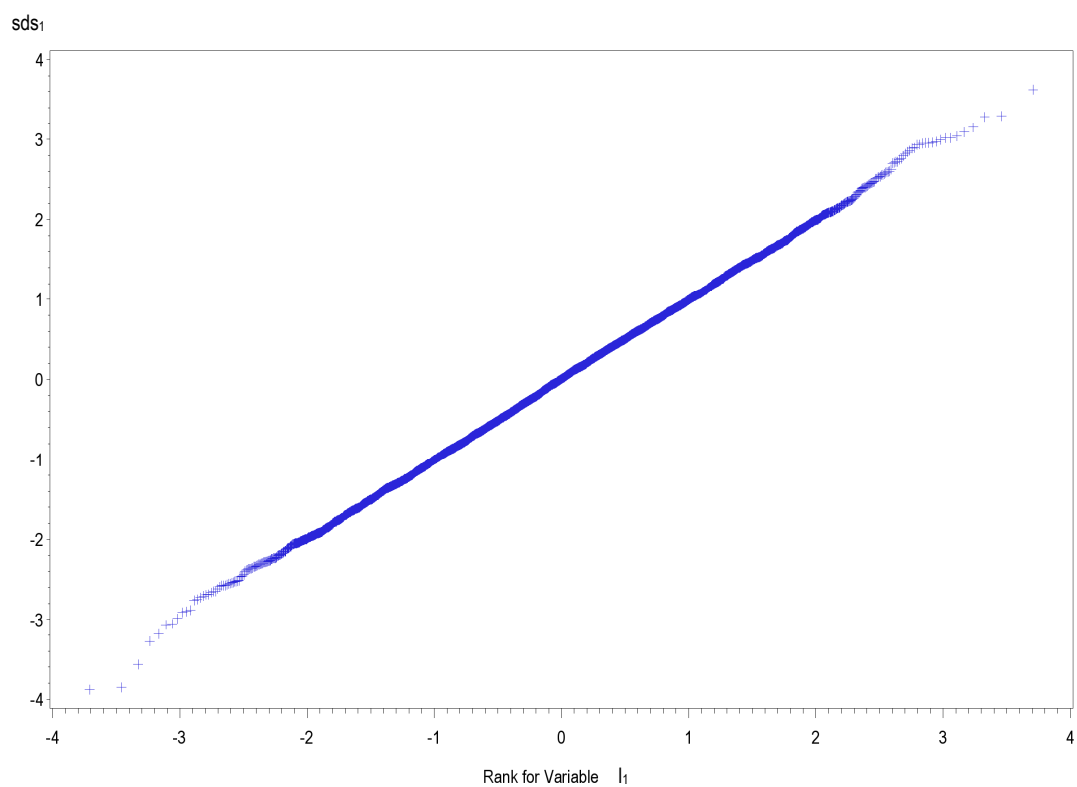


Figura A49-A72. Q-Qplot per l_1 - l_4 e sds_1 - sds_4 relativi agli scenari con 42 variabili per 75, 600 e 4800 soggetti

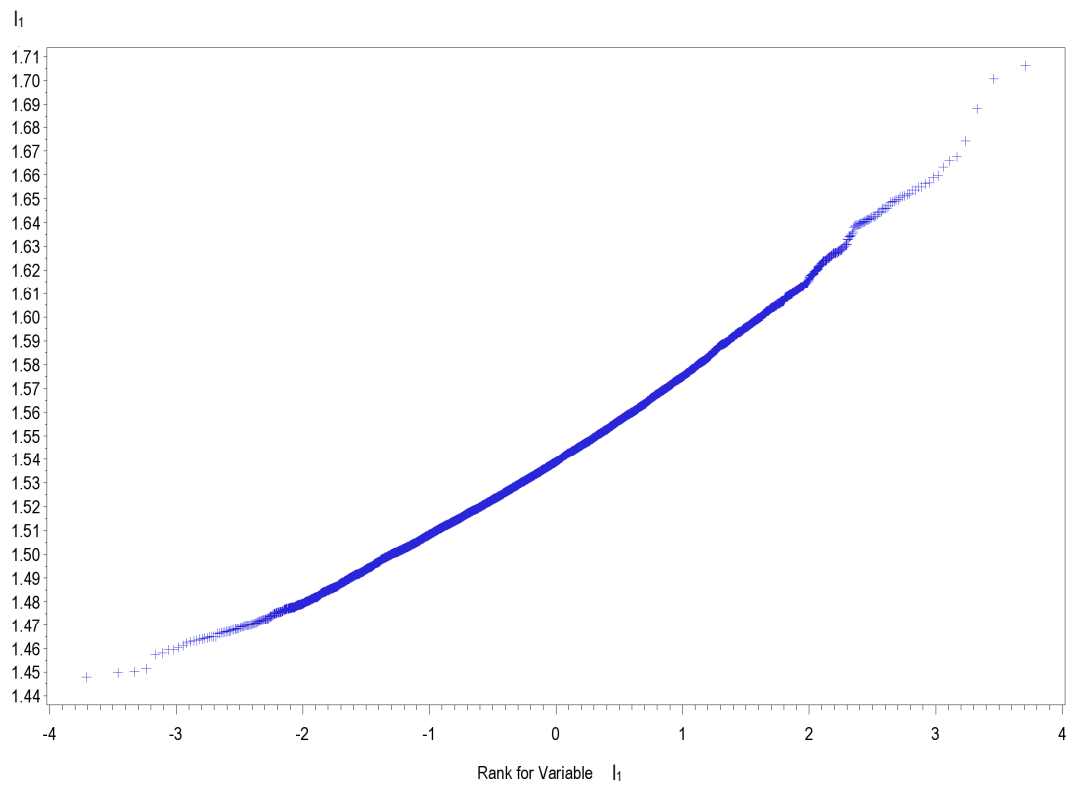
nsub=75 nvar=42



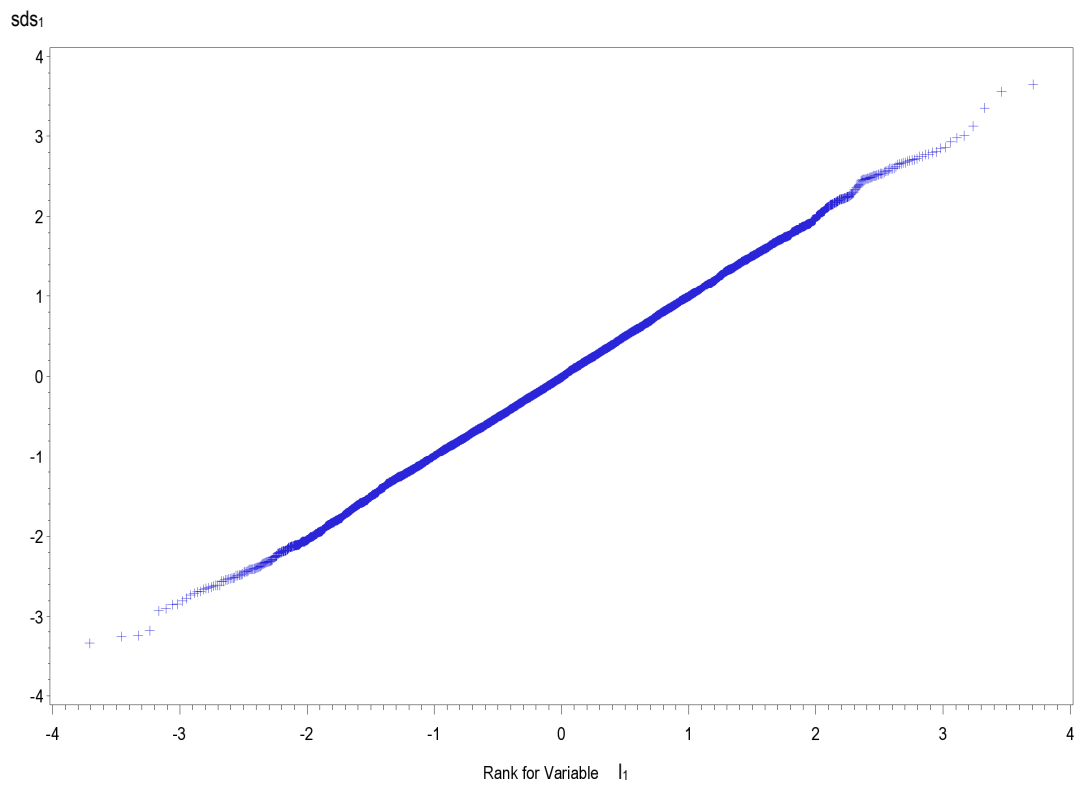
nsub=75 nvar=42



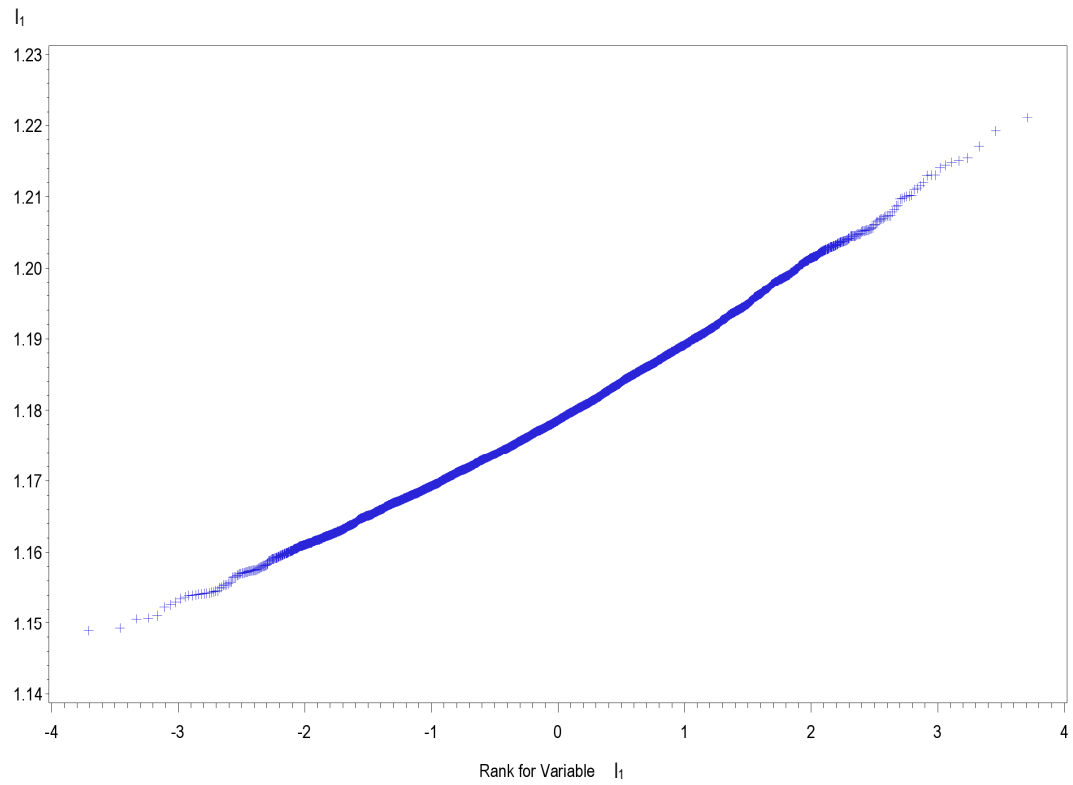
nsub=600 nvar=42



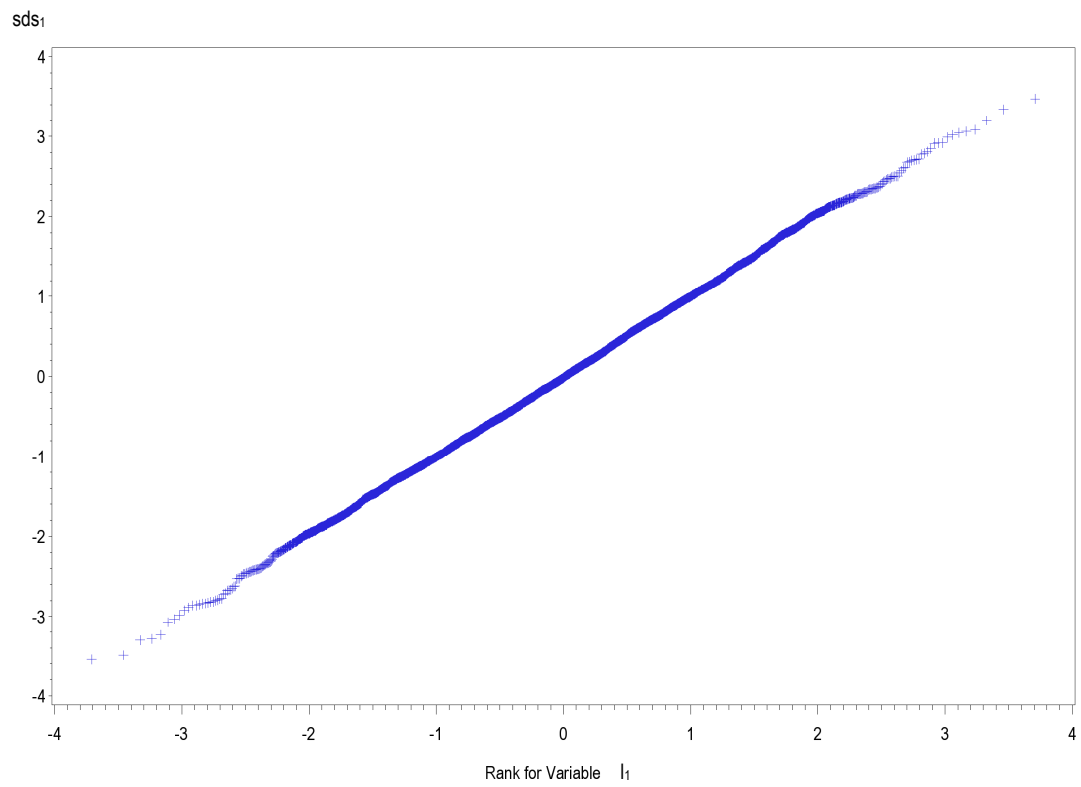
nsub=600 nvar=42



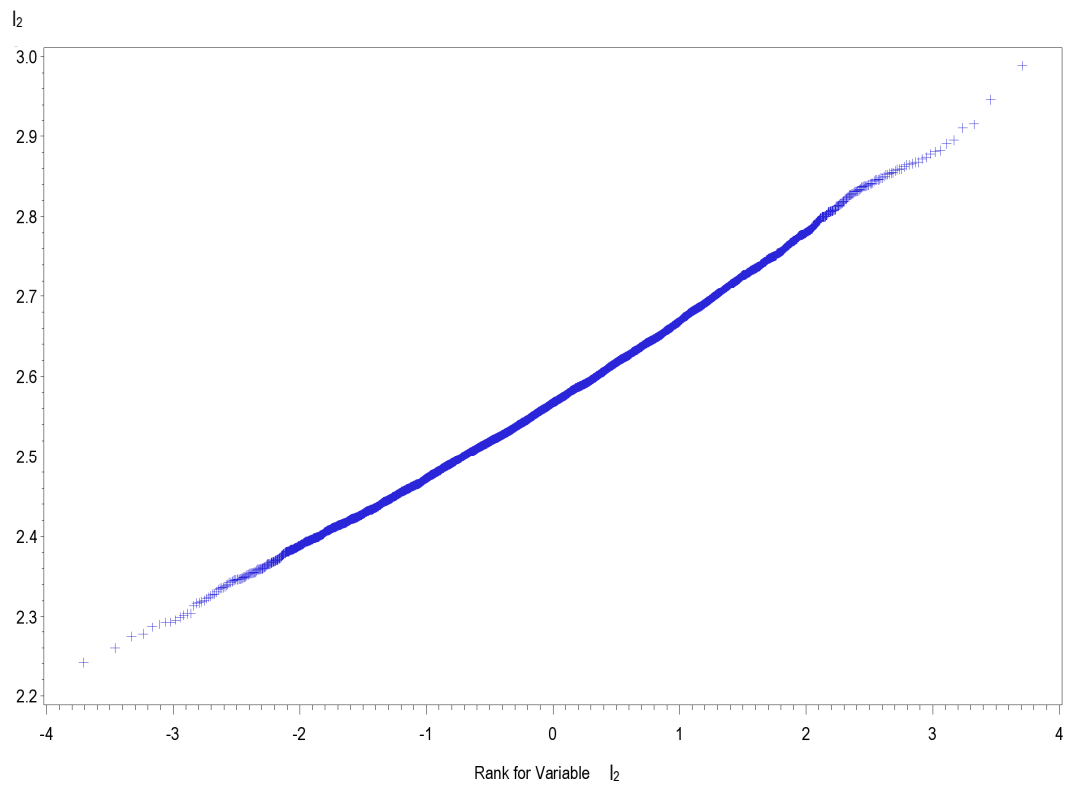
nsub=4800 nvar=42



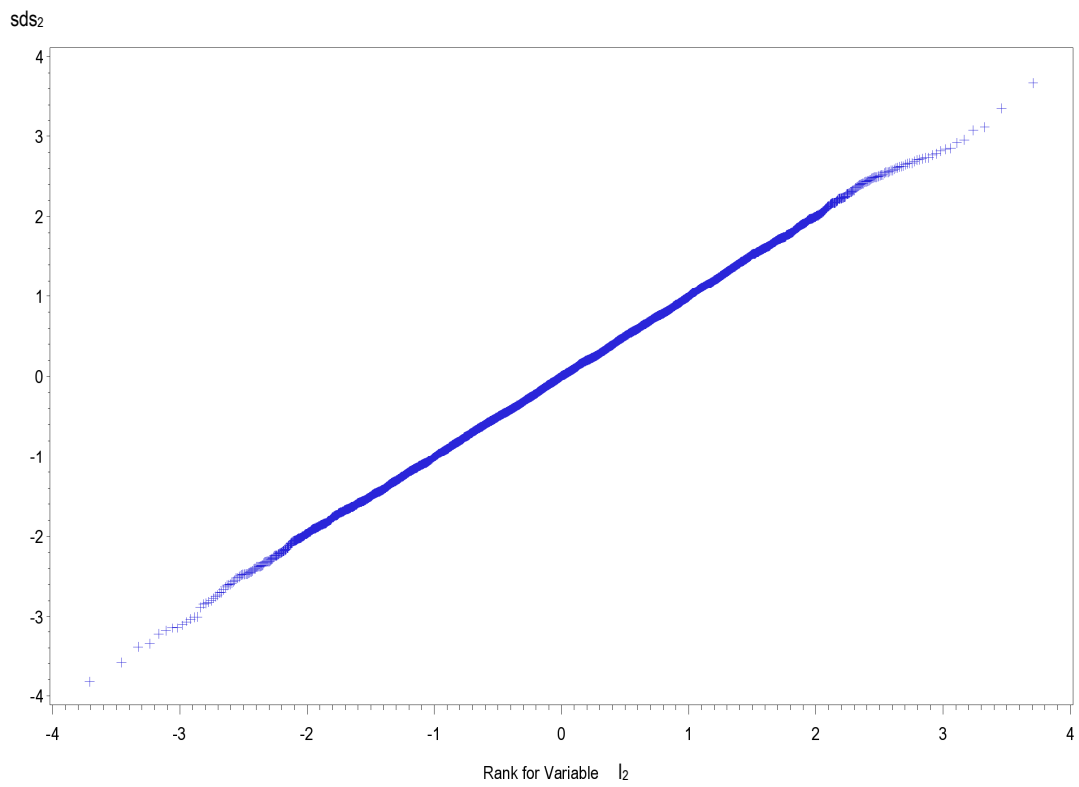
nsub=4800 nvar=42



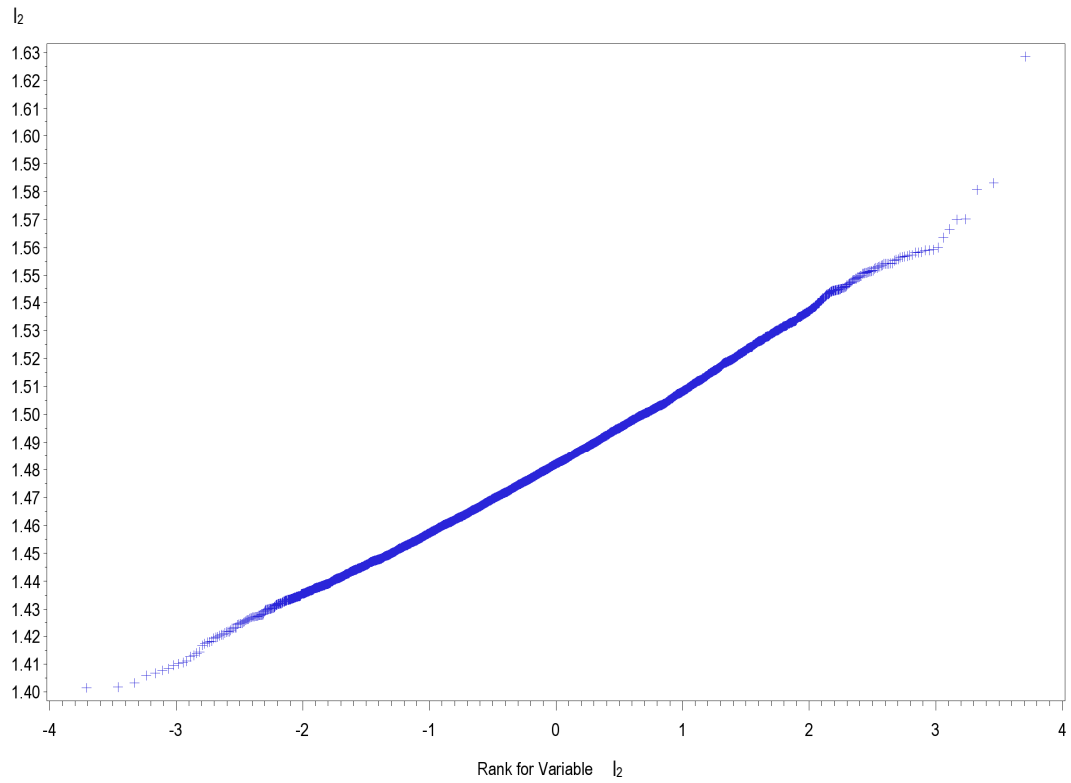
nsub=75 nvar=42



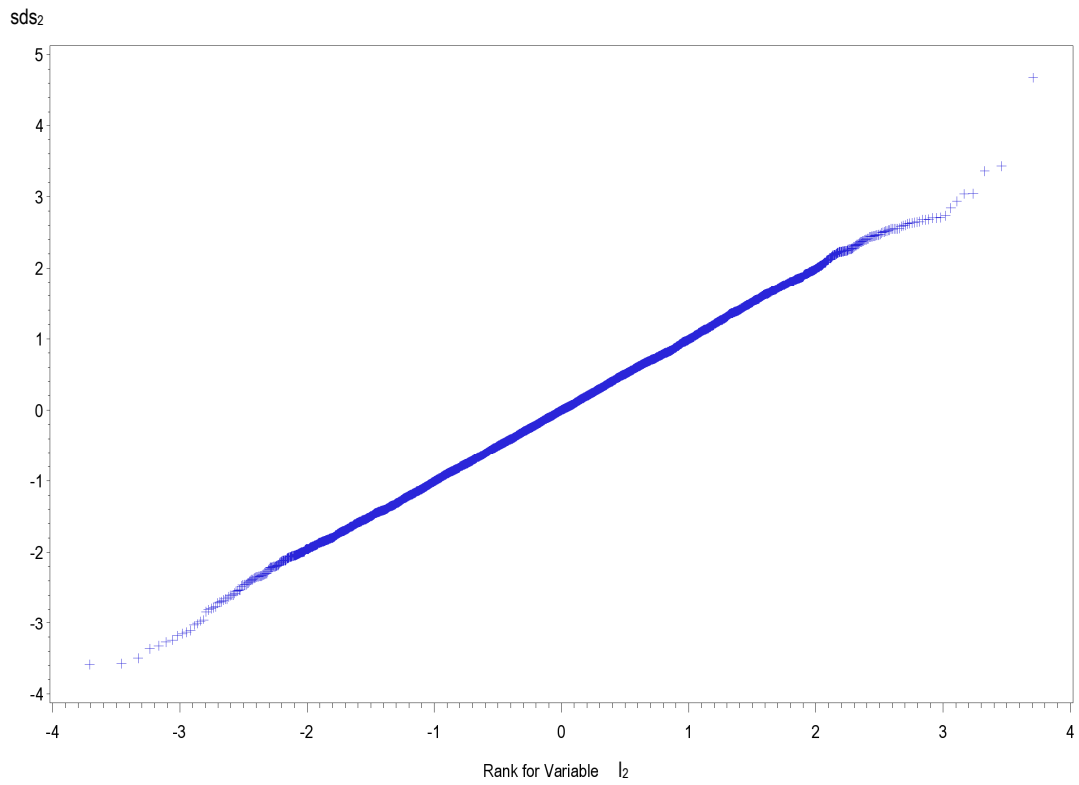
nsub=75 nvar=42



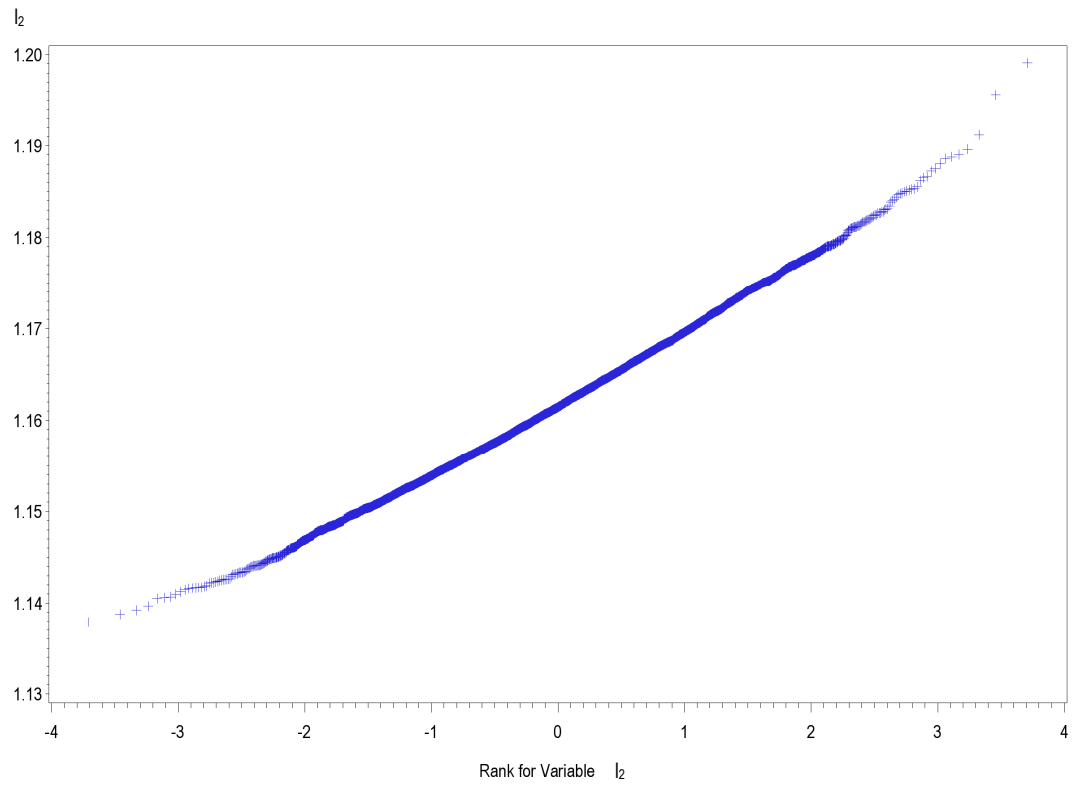
nsub=600 nvar=42



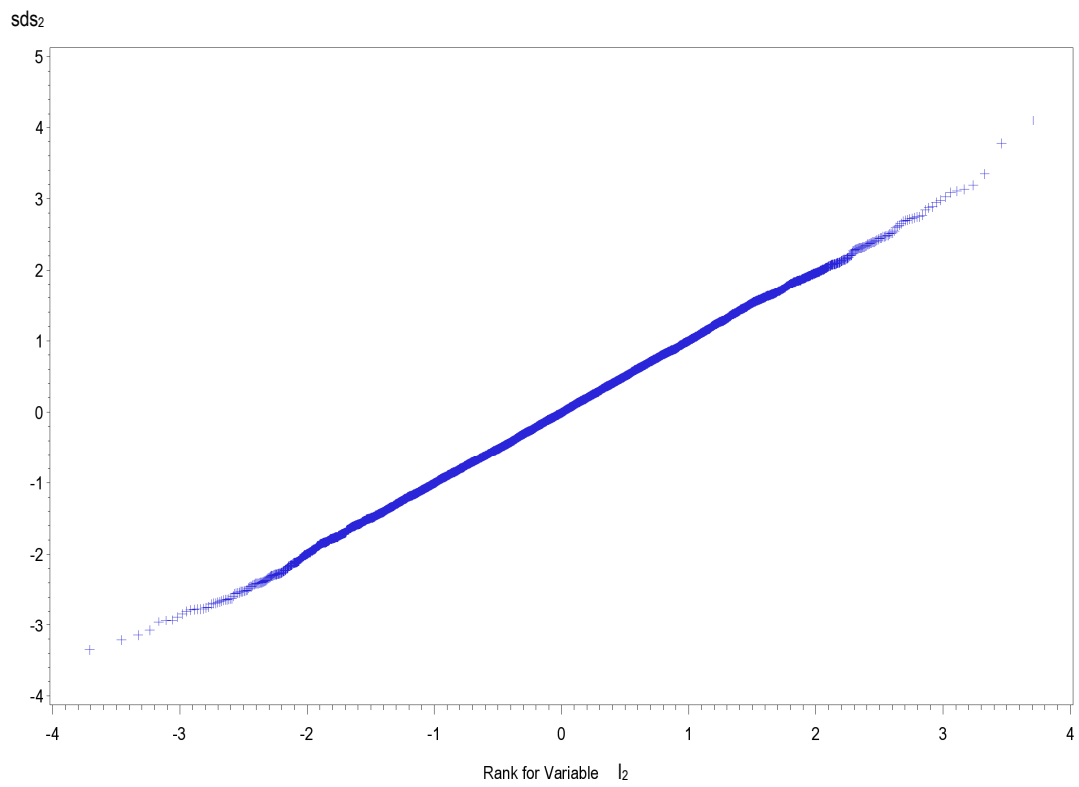
nsub=600 nvar=42



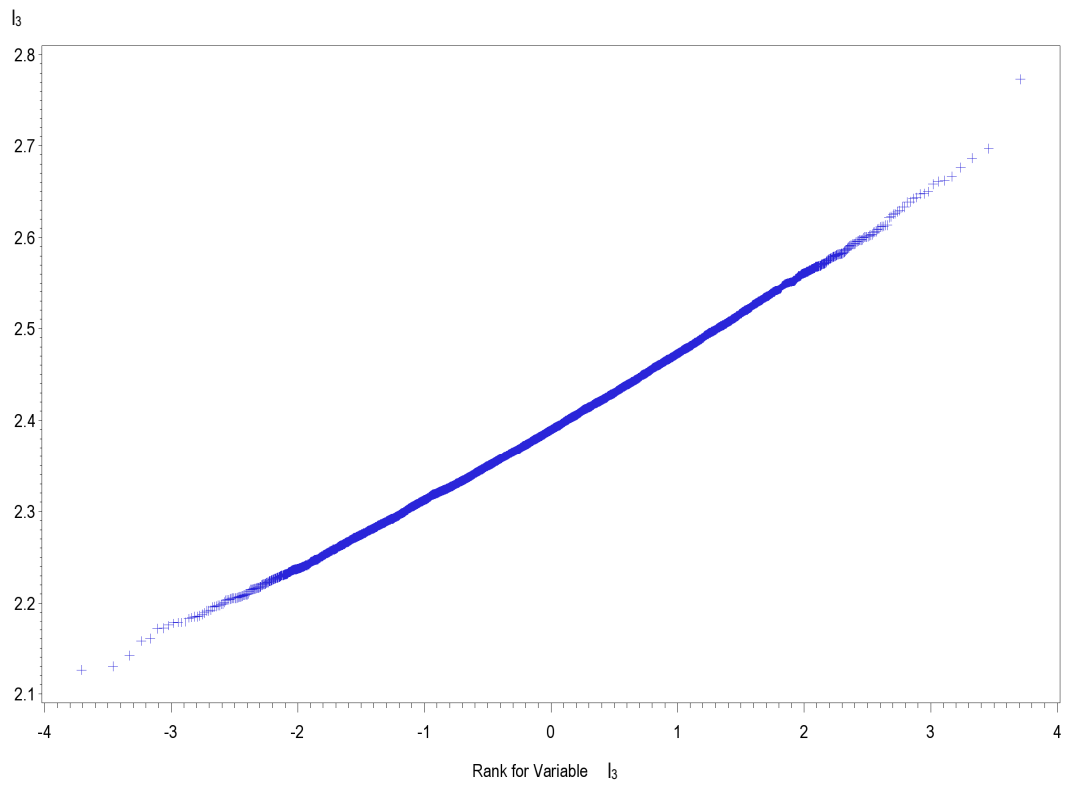
nsub=4800 nvar=42



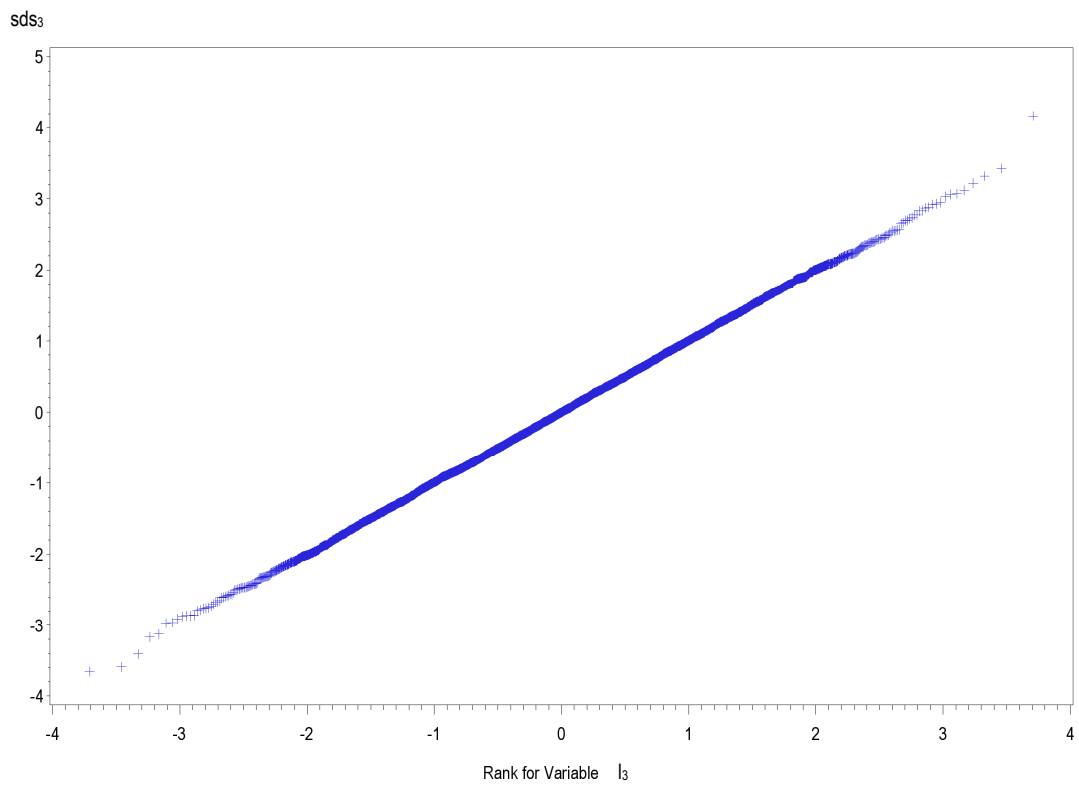
nsub=4800 nvar=42



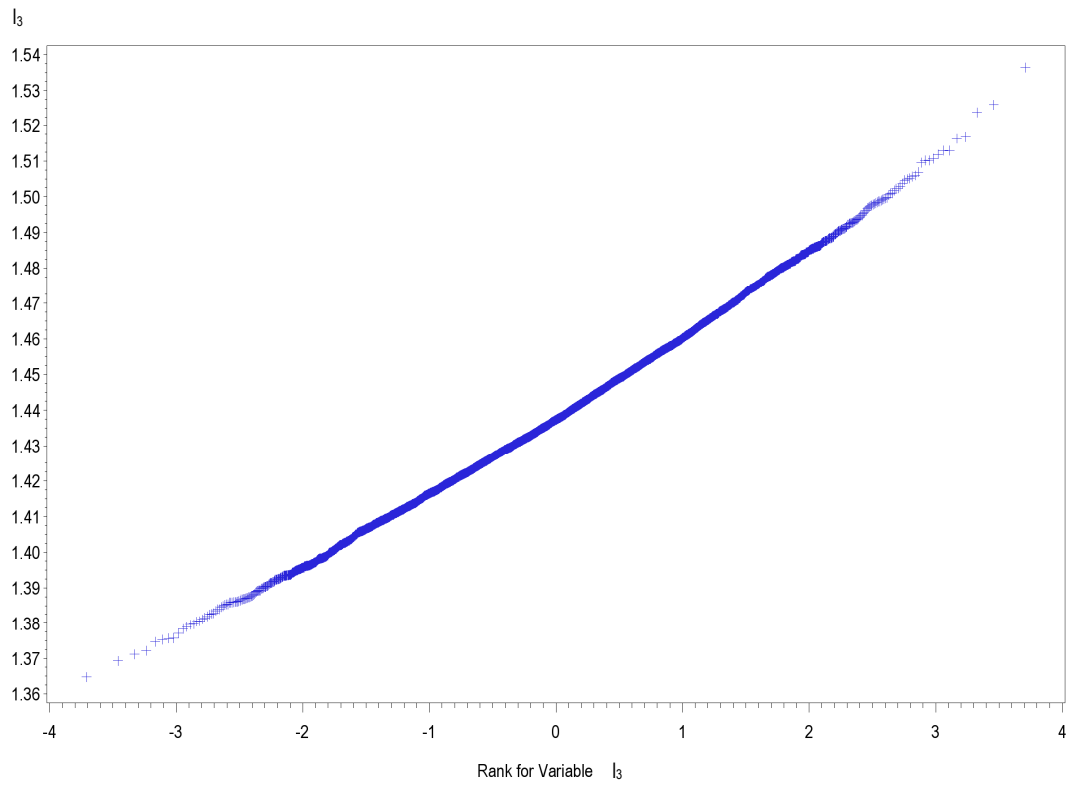
nsub=75 nvar=42



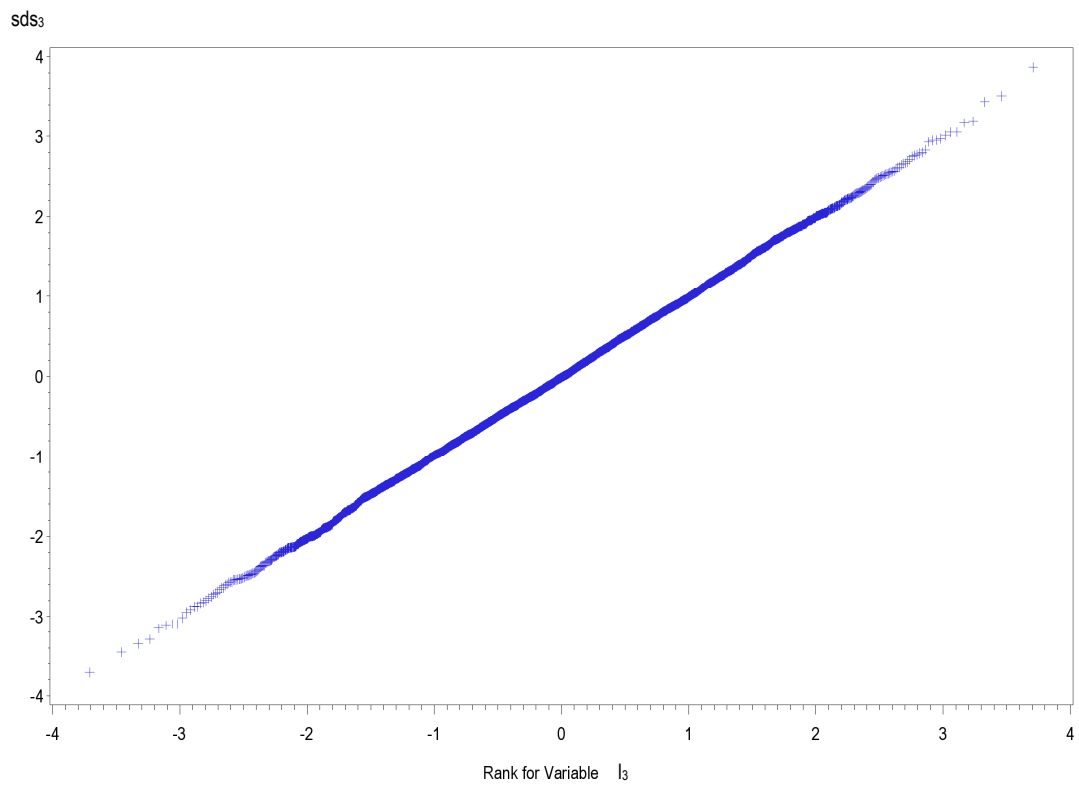
nsub=75 nvar=42



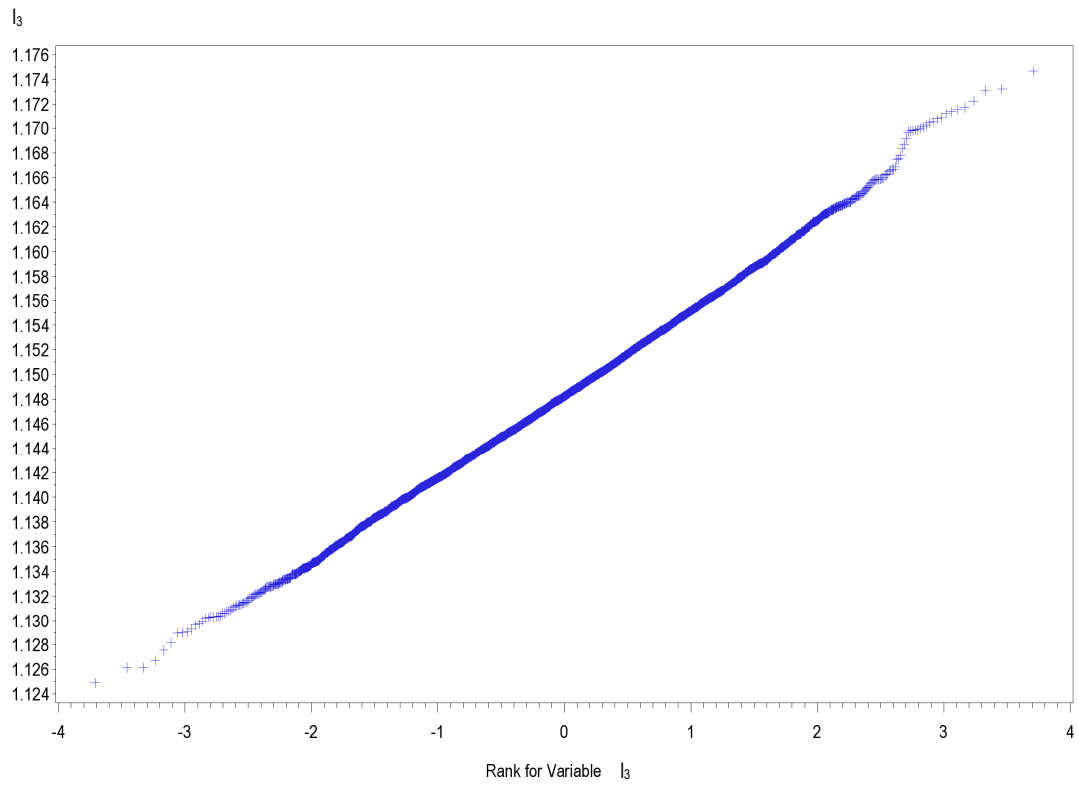
nsub=600 nvar=42



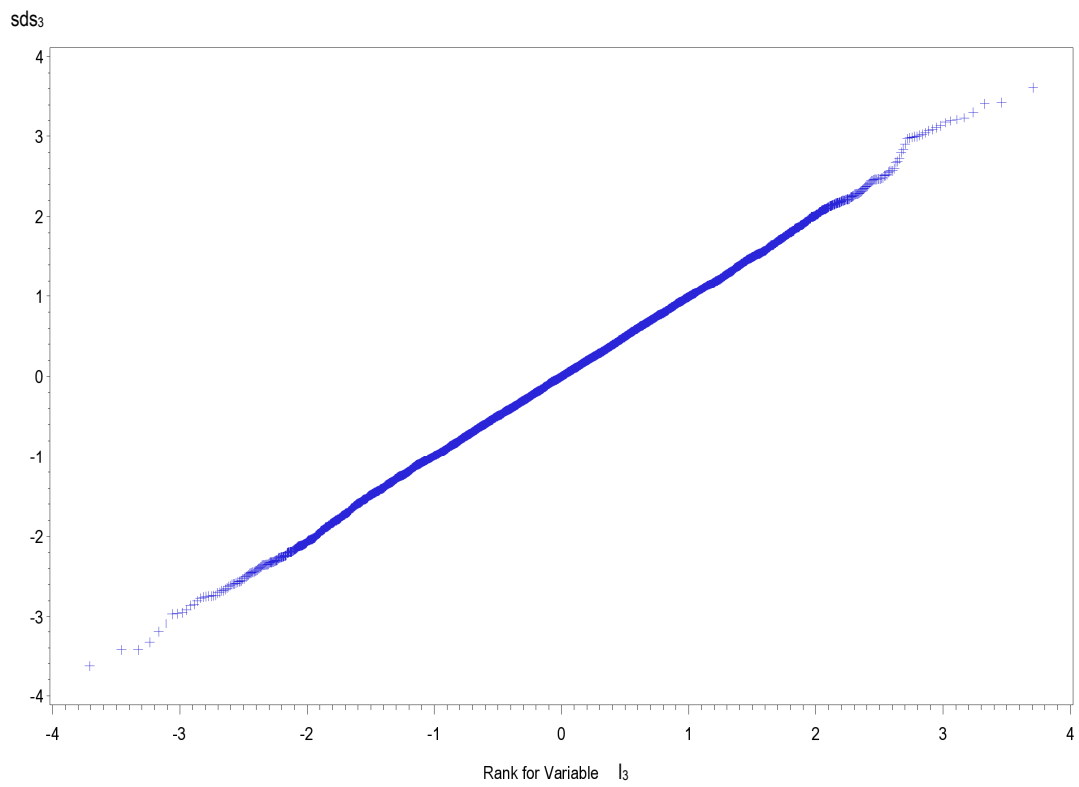
nsub=600 nvar=42



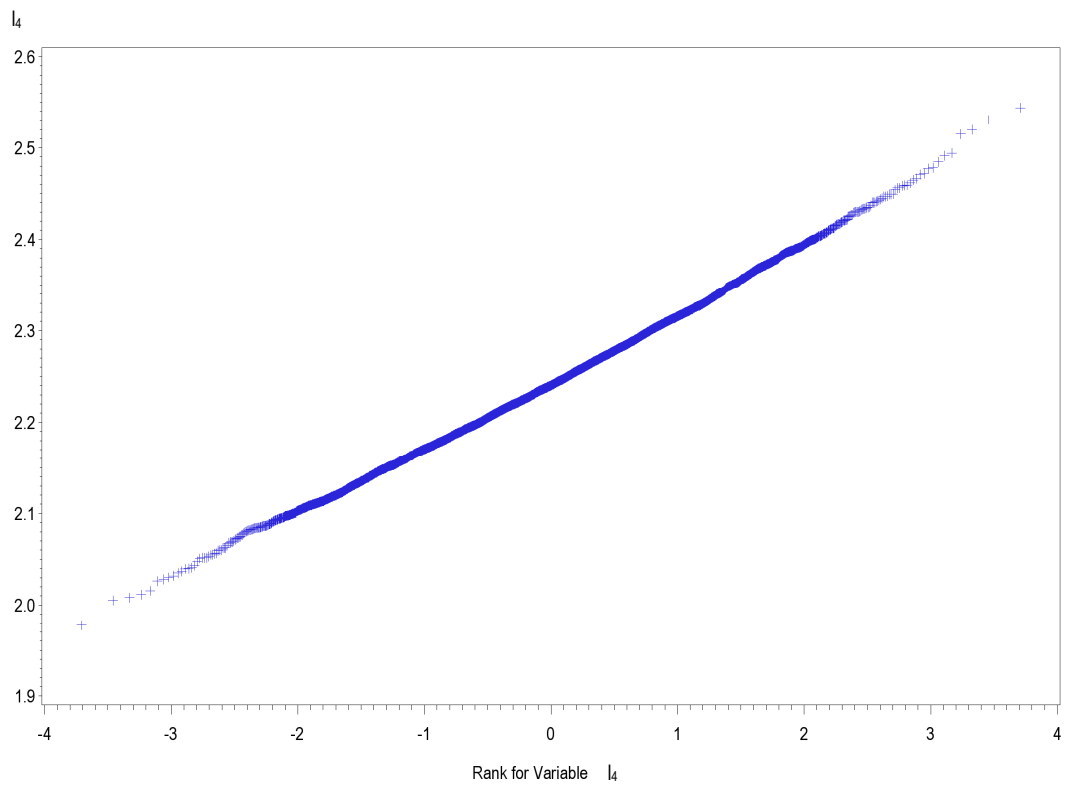
nsub=4800 nvar=42



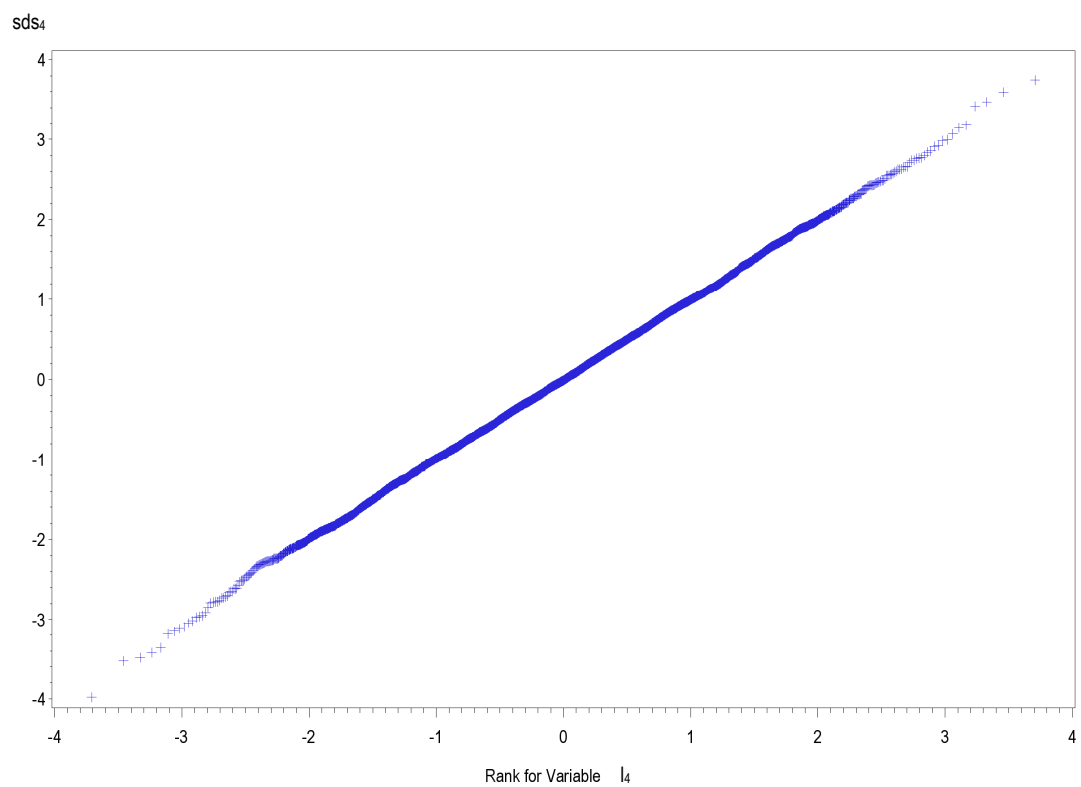
nsub=4800 nvar=42



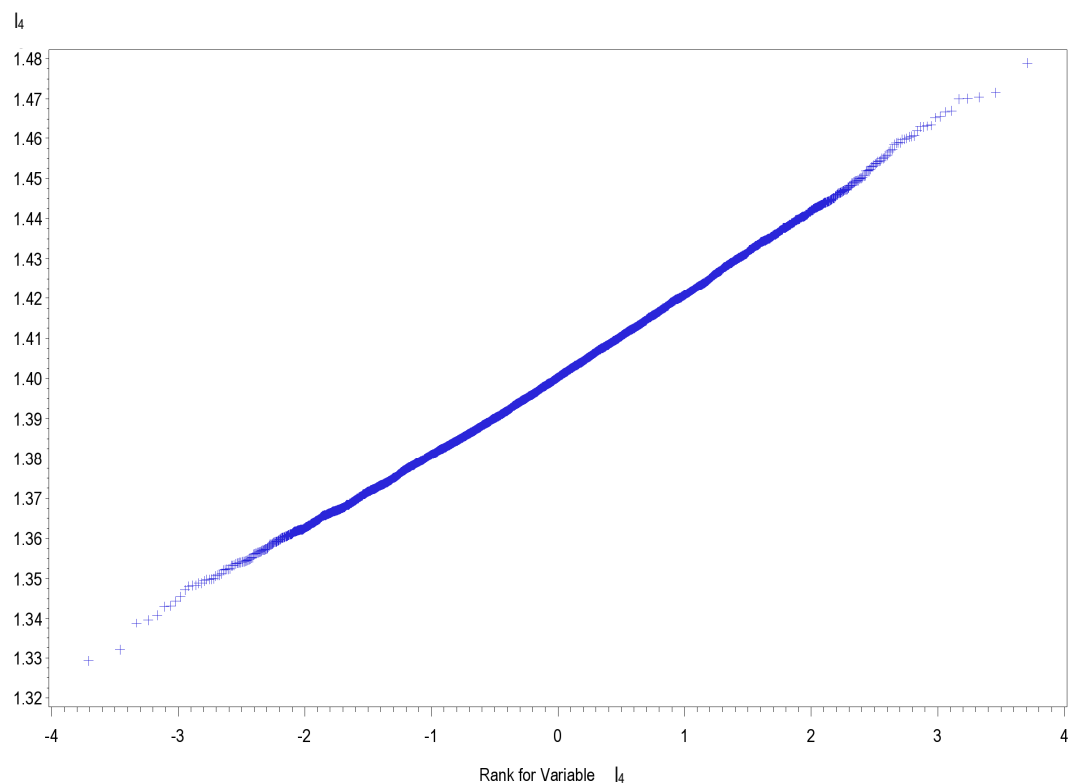
nsub=75 nvar=42



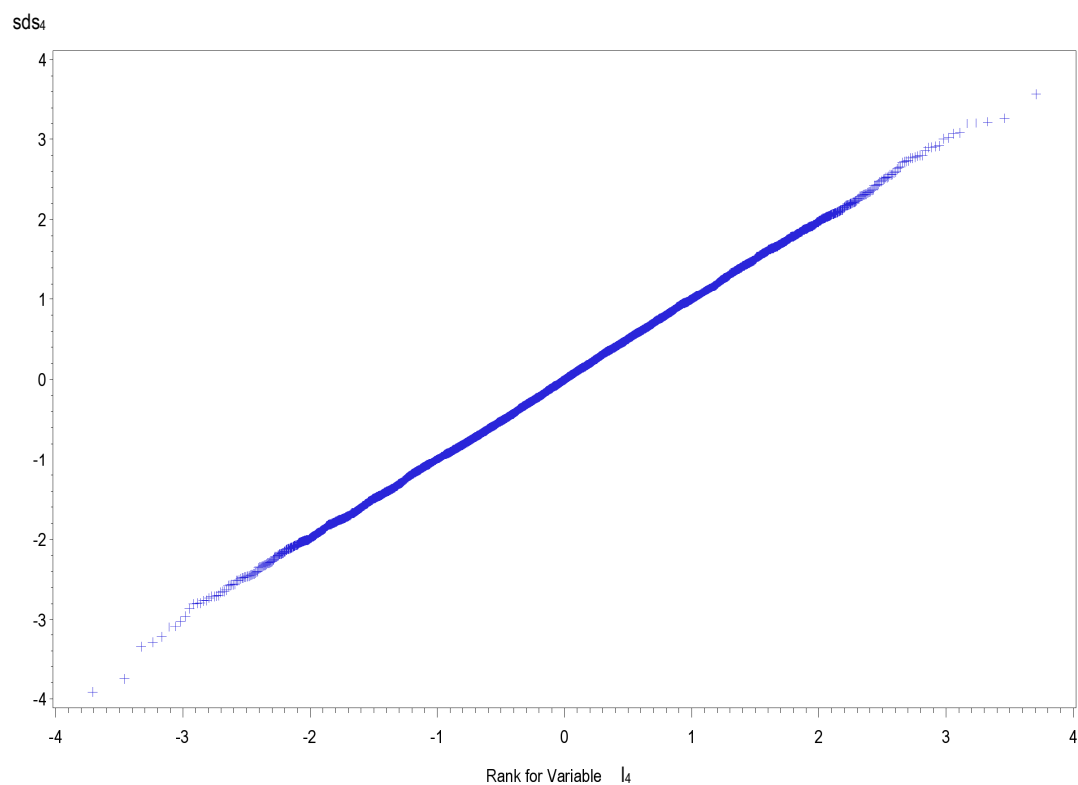
nsub=75 nvar=42



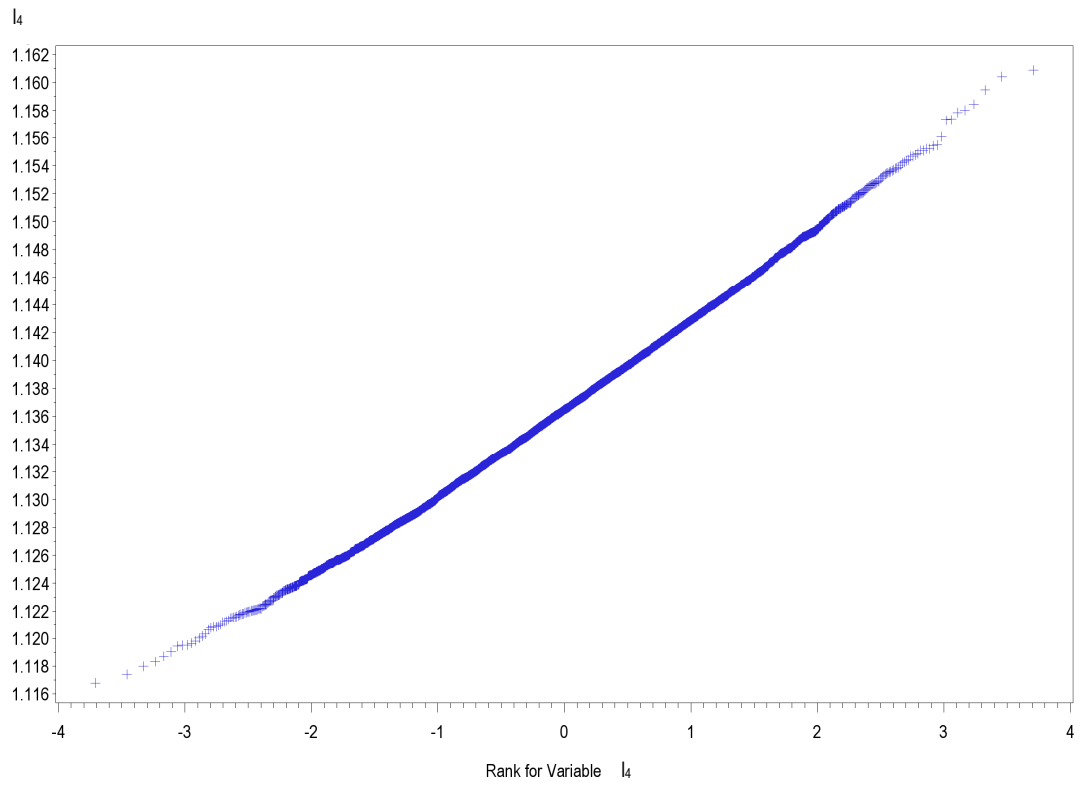
nsub=600 nvar=42



nsub=600 nvar=42



nsub=4800 nvar=42



nsub=4800 nvar=42

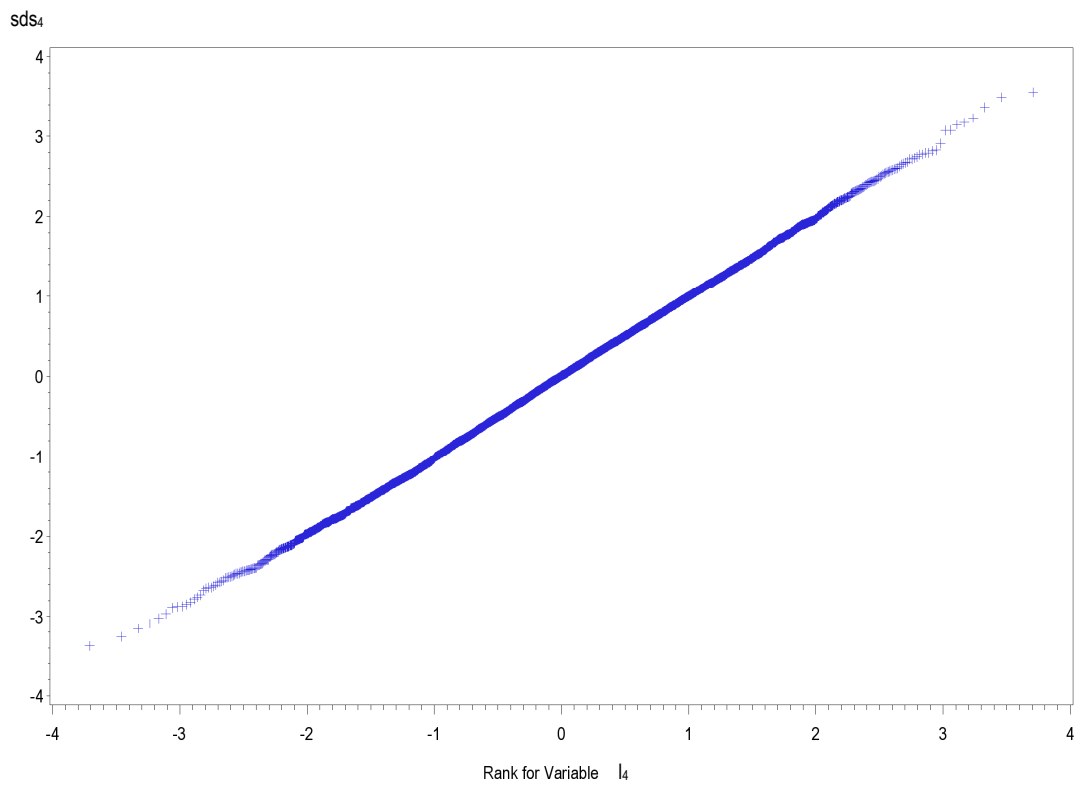
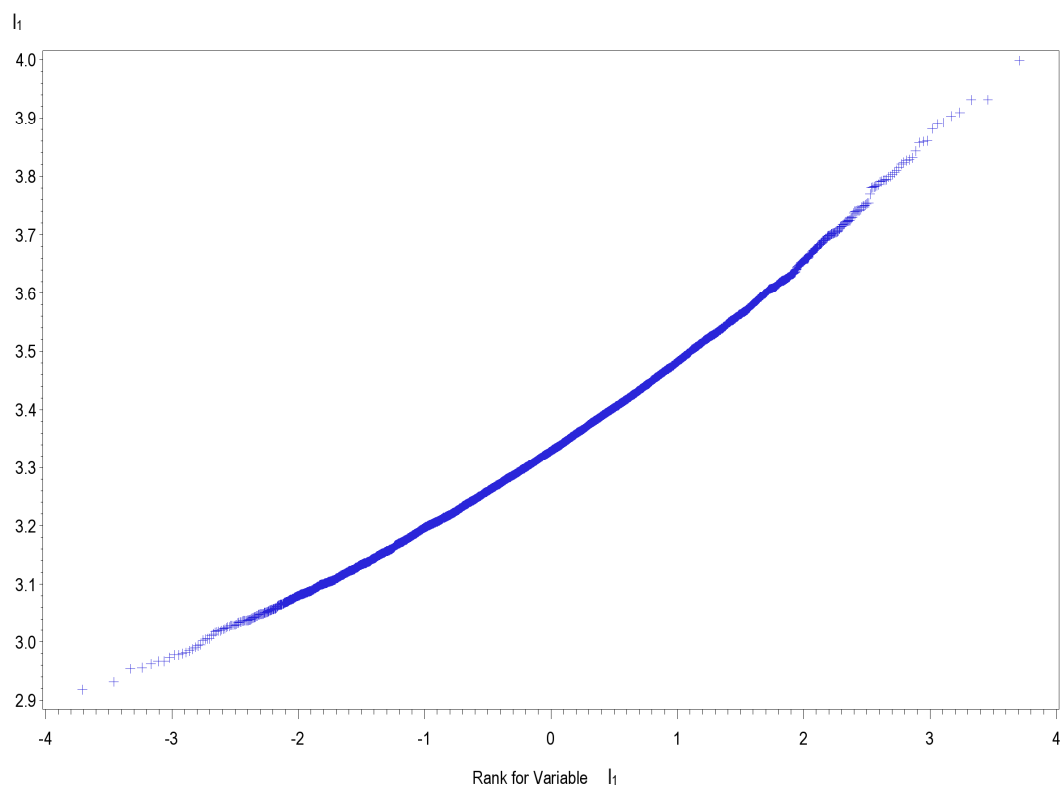
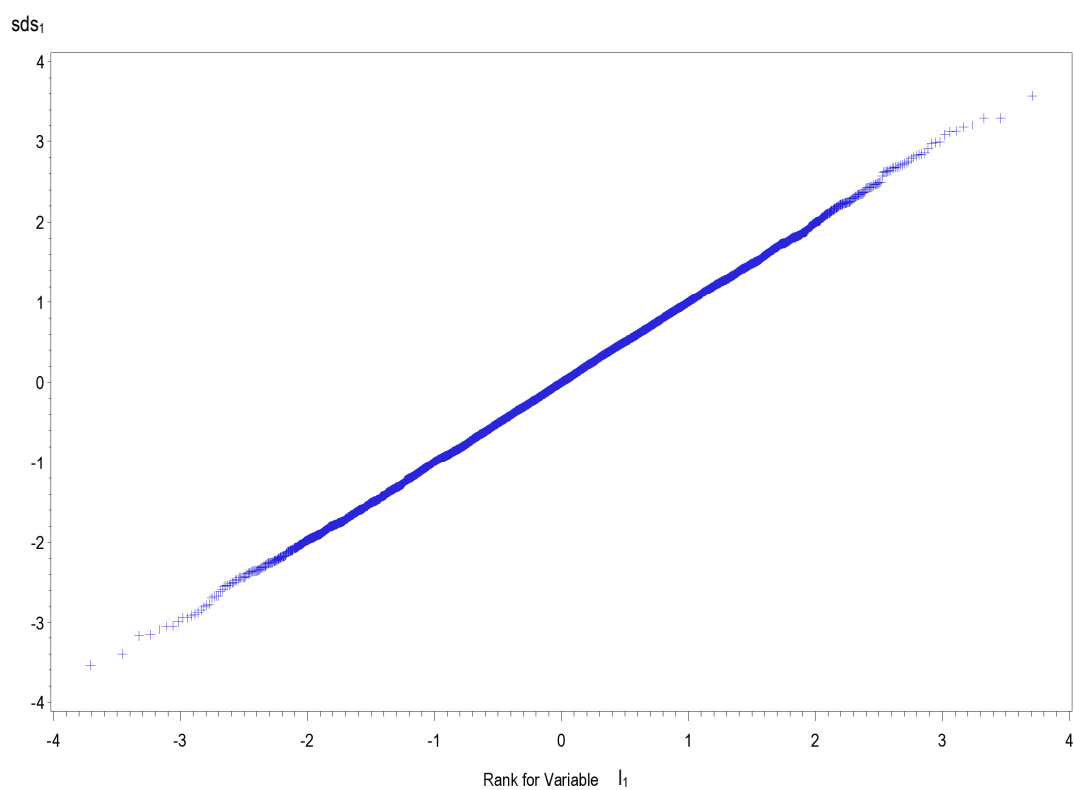


Figura A73-A96. Q-Qplot per l_1 - l_4 e sds_1 - sds_4 relativi agli scenari con 60 variabili per 75, 600 e 4800 soggetti

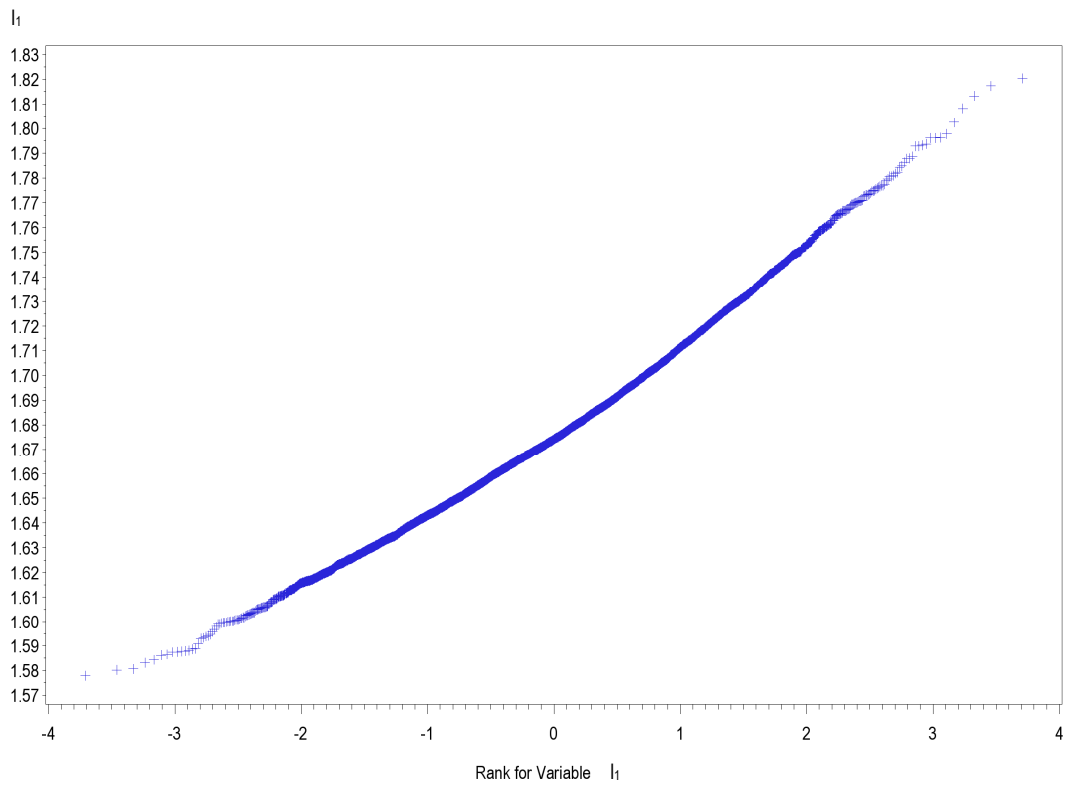
nsub=75 nvar=60



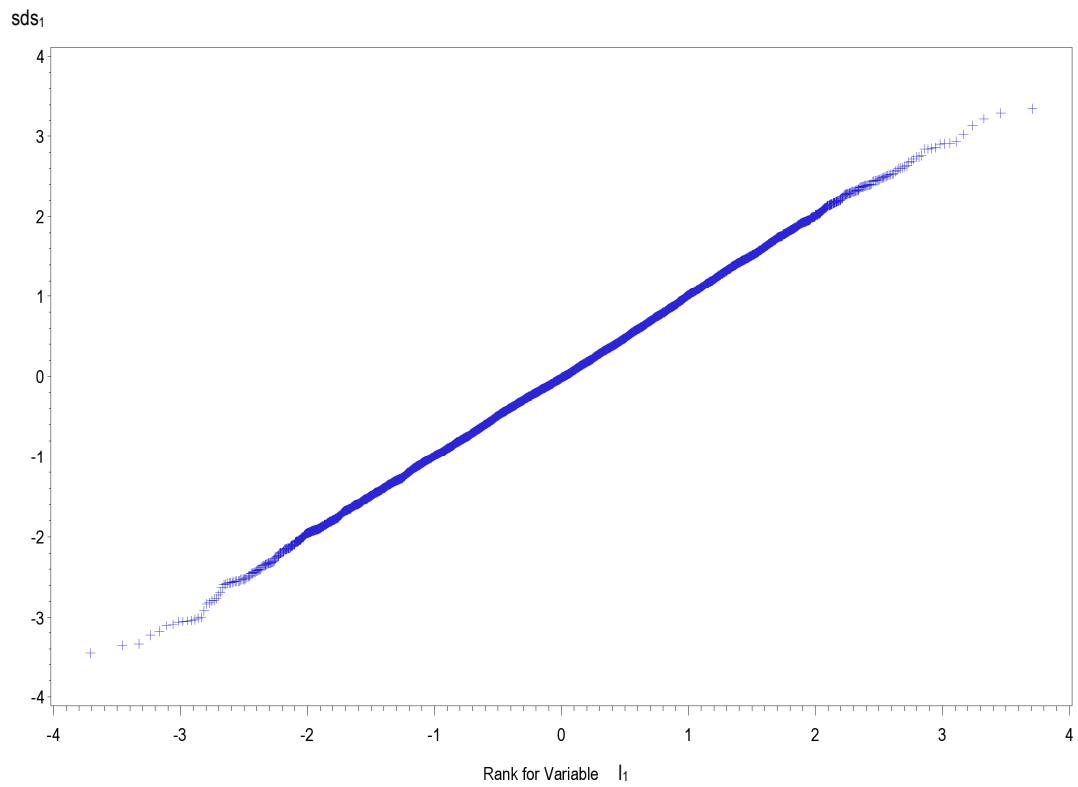
nsub=75 nvar=60



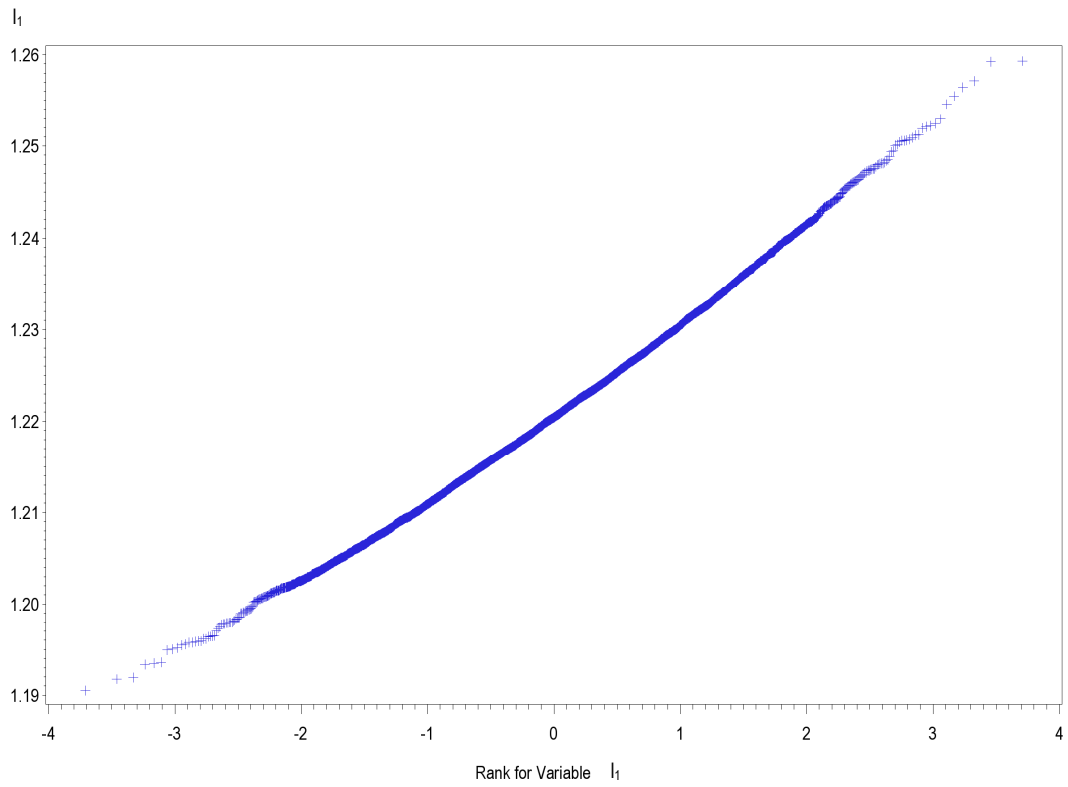
nsub=600 nvar=60



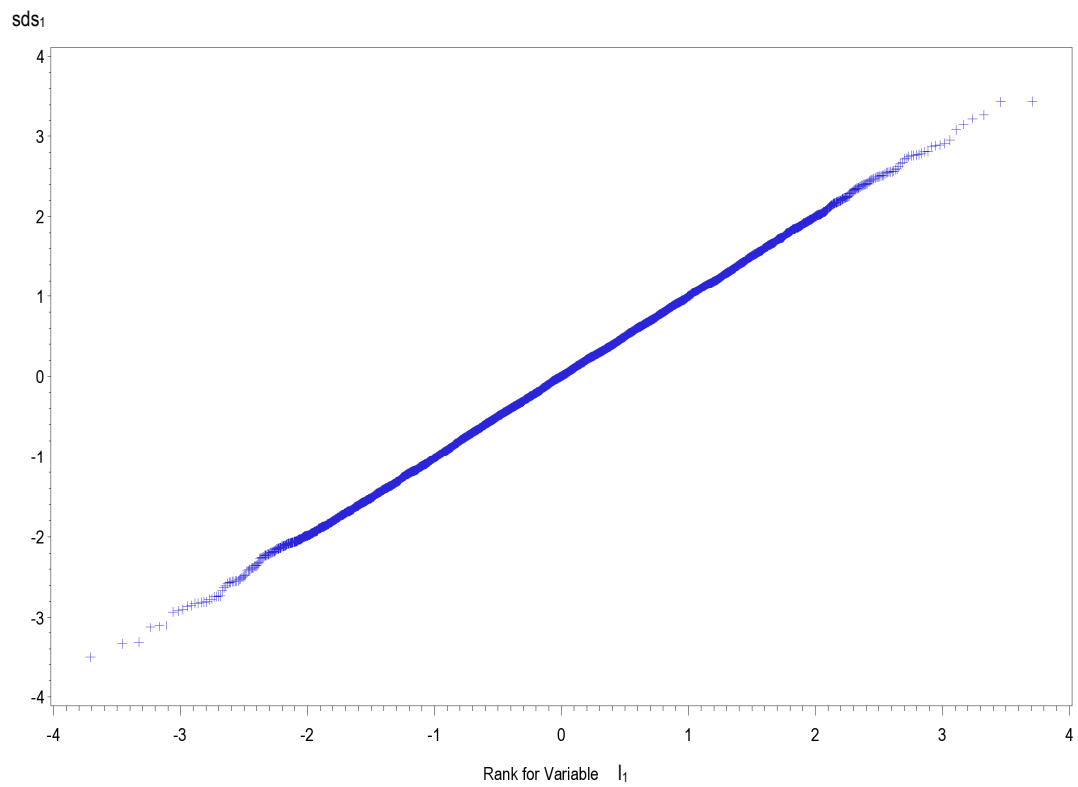
nsub=600 nvar=60



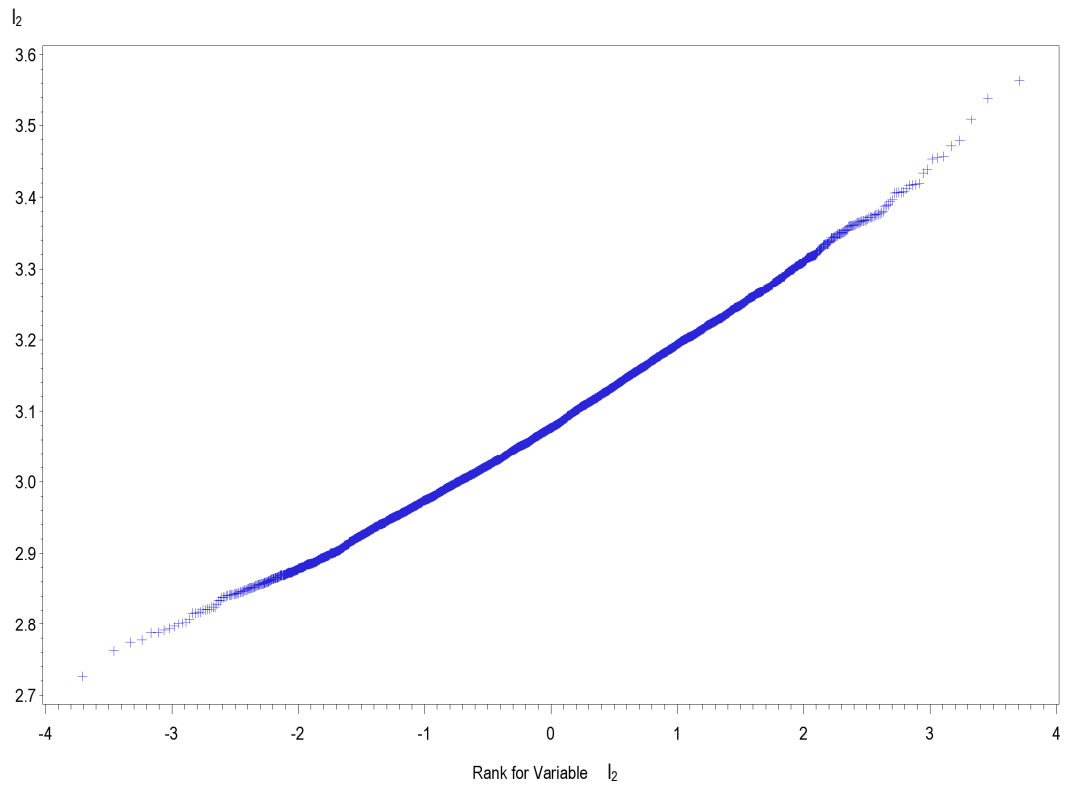
nsub=4800 nvar=60



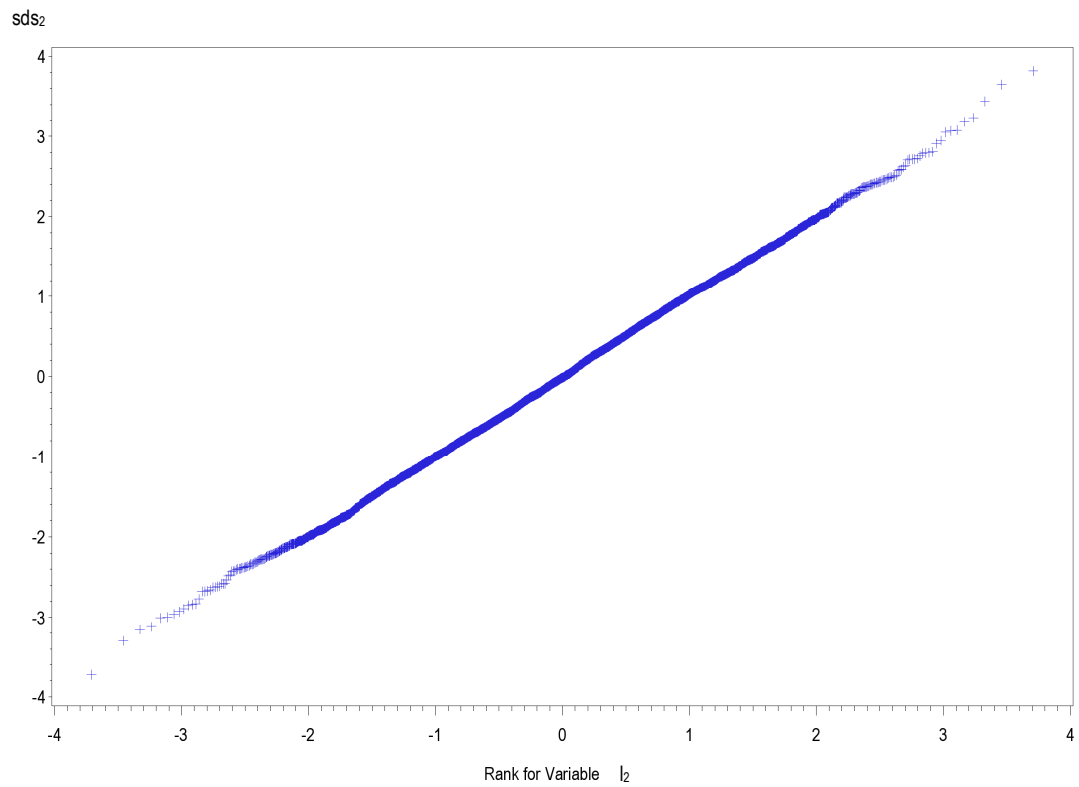
nsub=4800 nvar=60



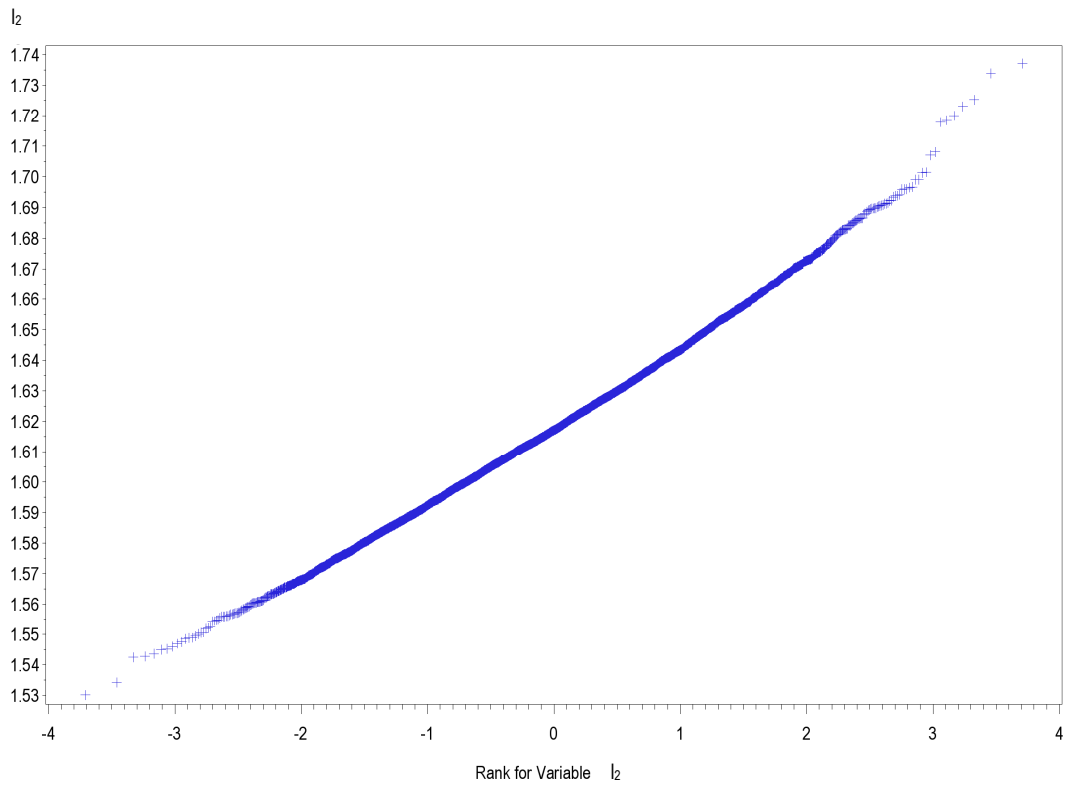
nsub=75 nvar=60



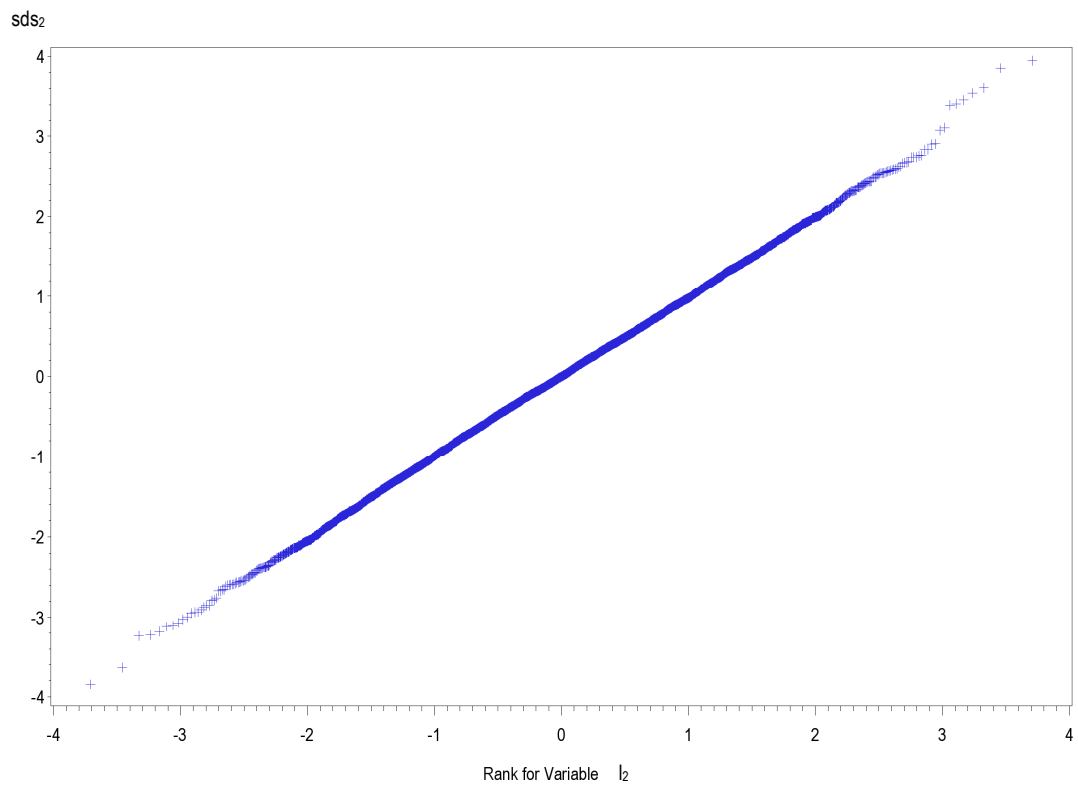
nsub=75 nvar=60



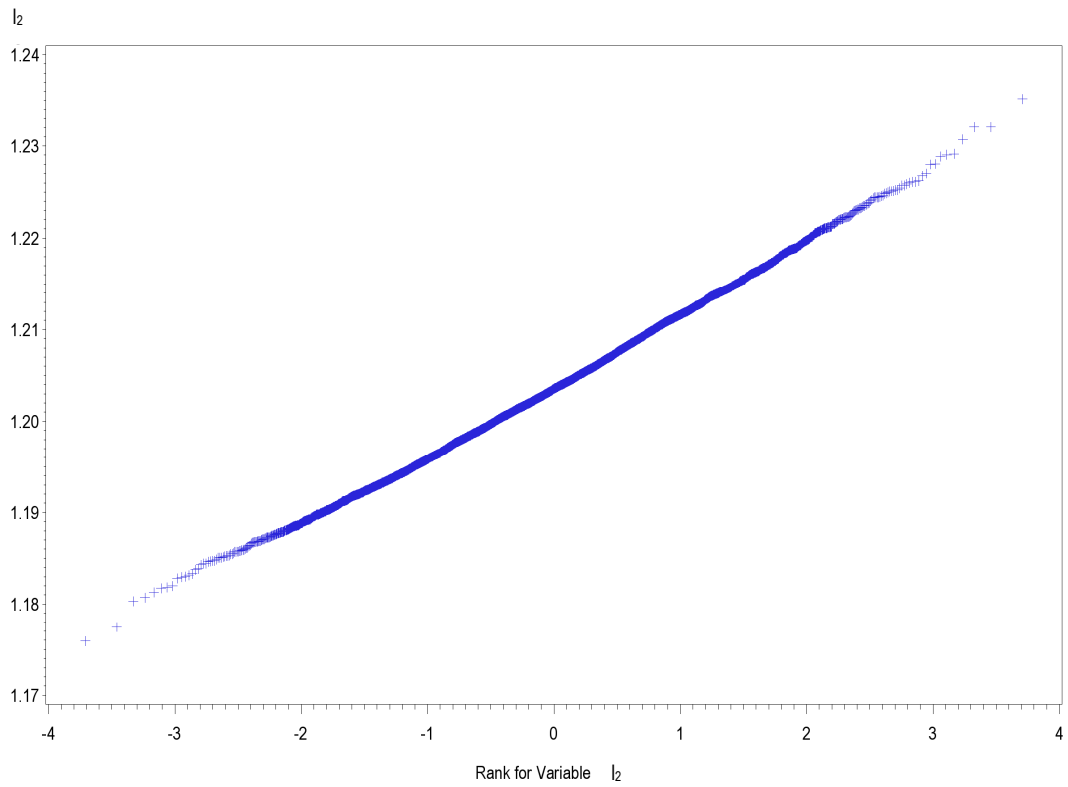
nsub=600 nvar=60



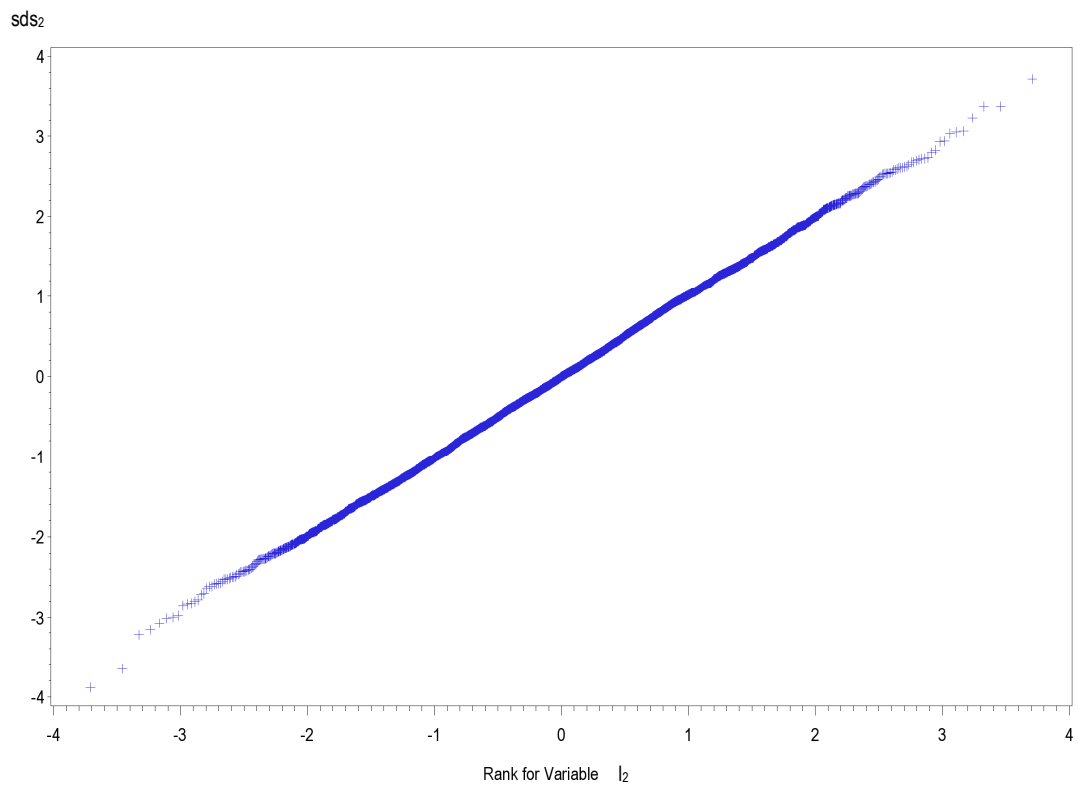
nsub=600 nvar=60



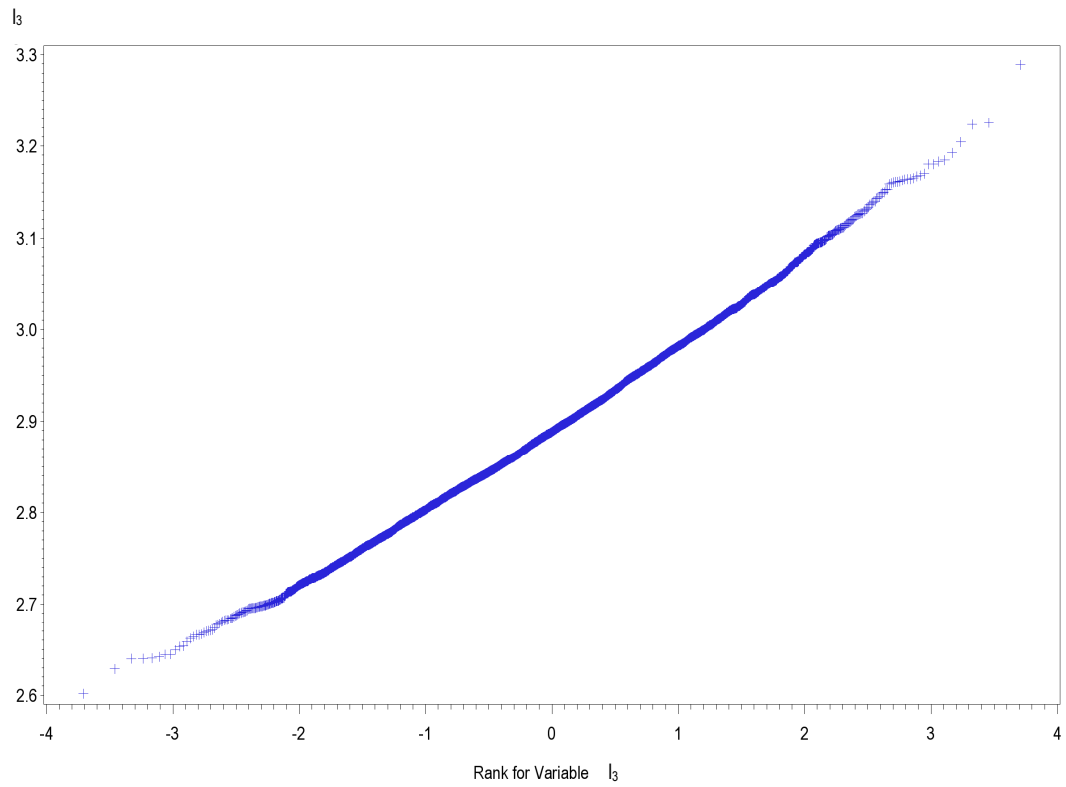
nsub=4800 nvar=60



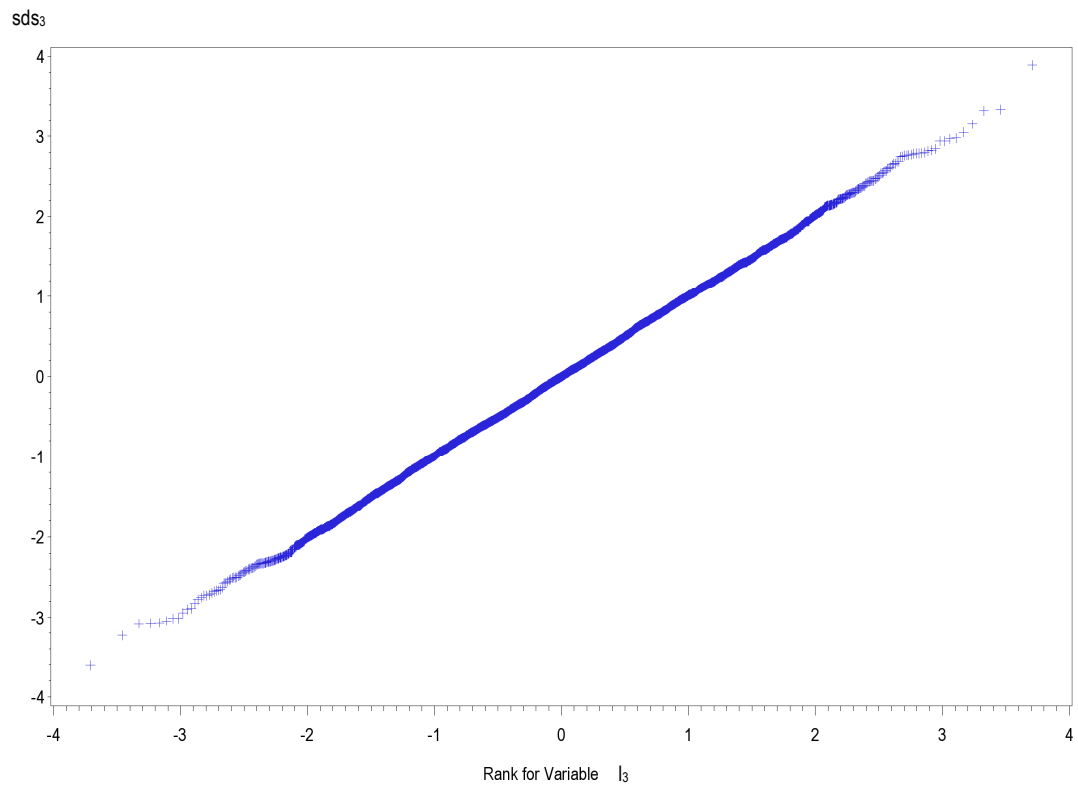
nsub=4800 nvar=60



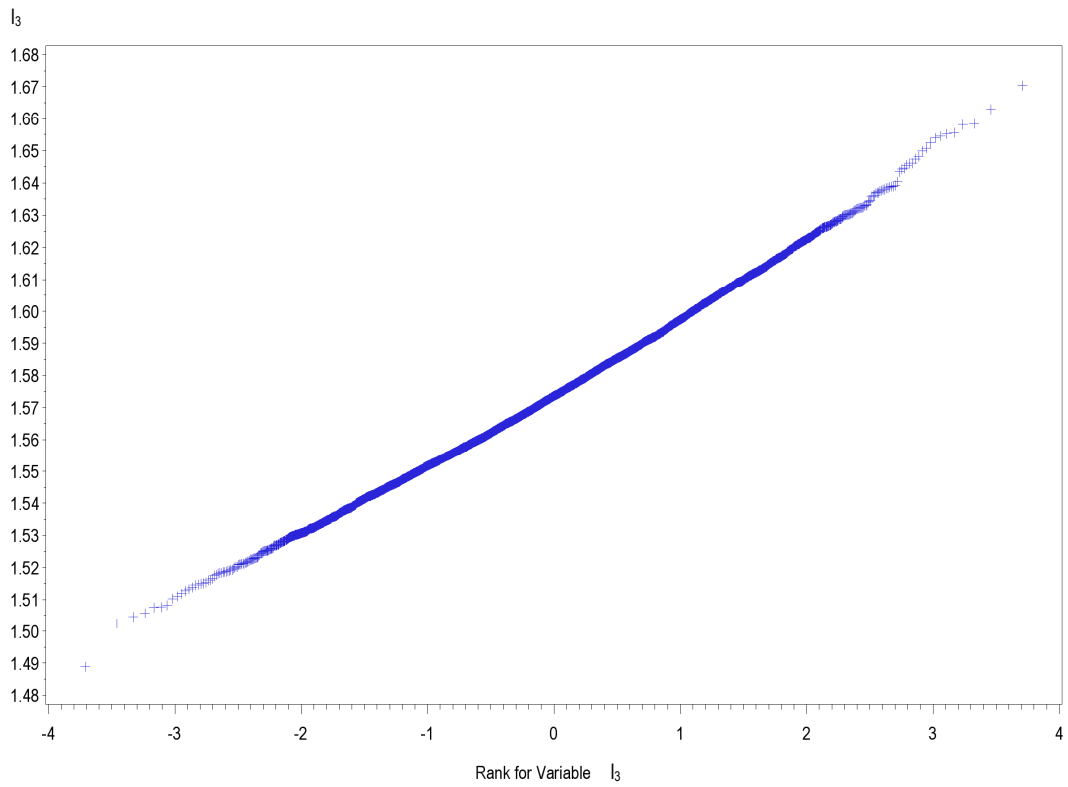
nsub=75 nvar=60



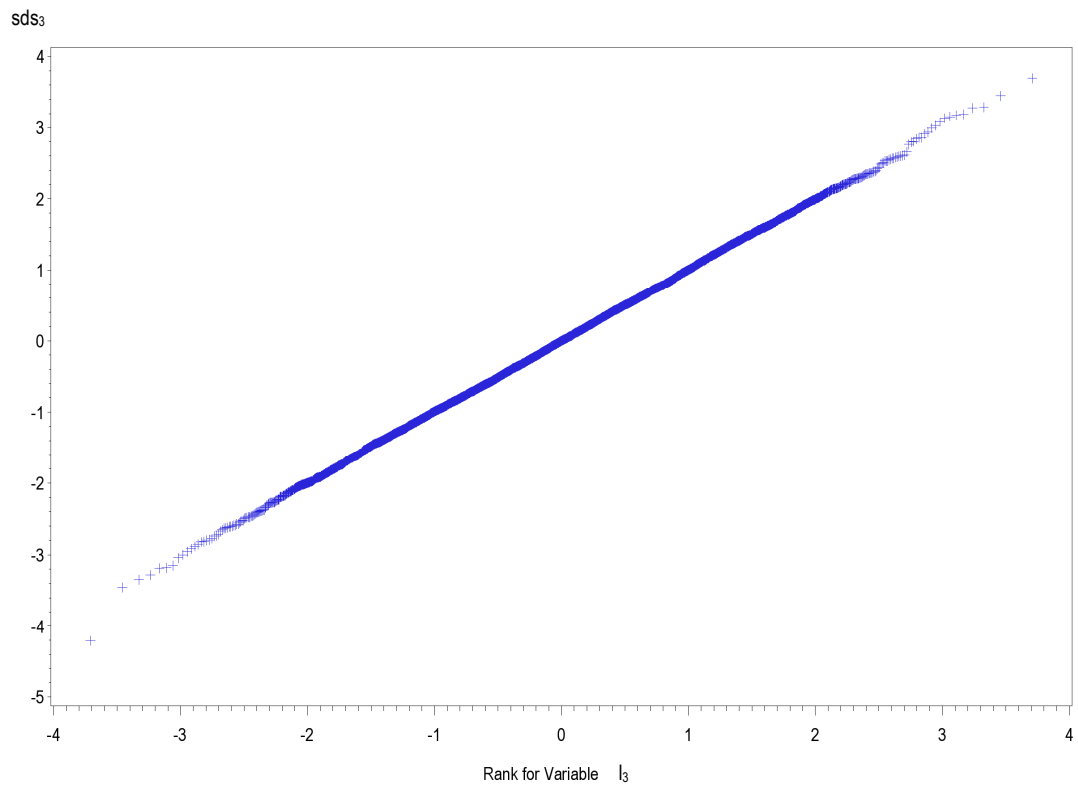
nsub=75 nvar=60



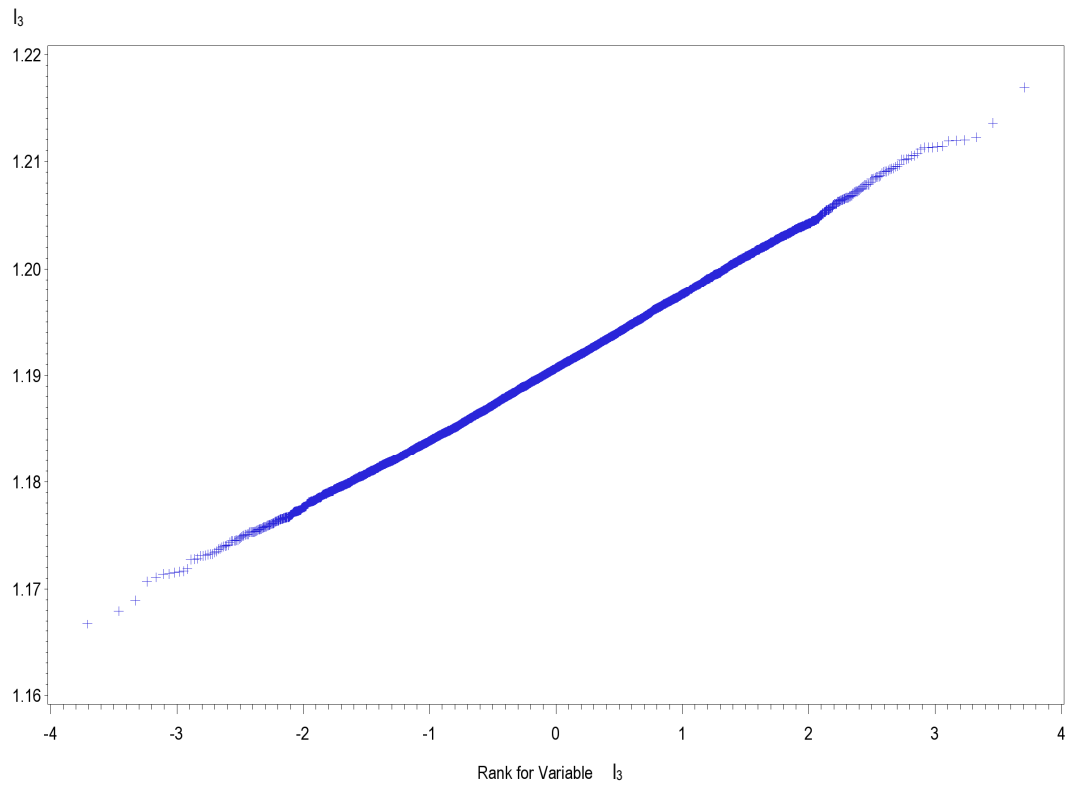
nsub=600 nvar=60



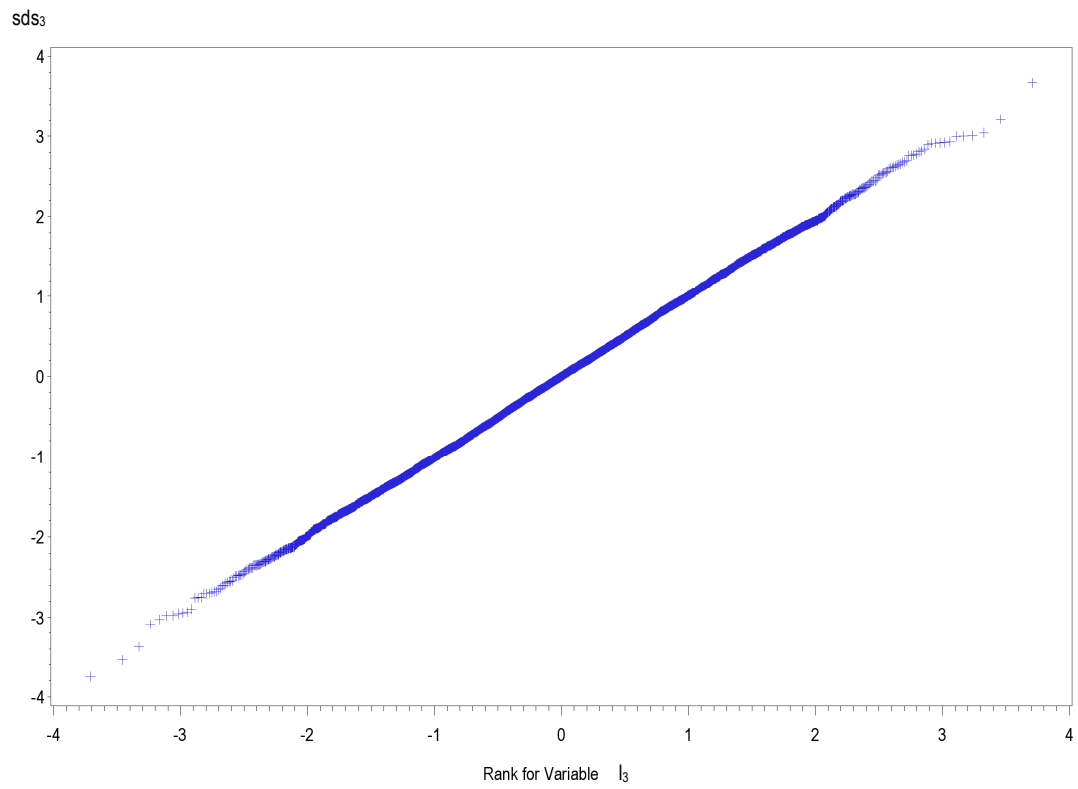
nsub=600 nvar=60



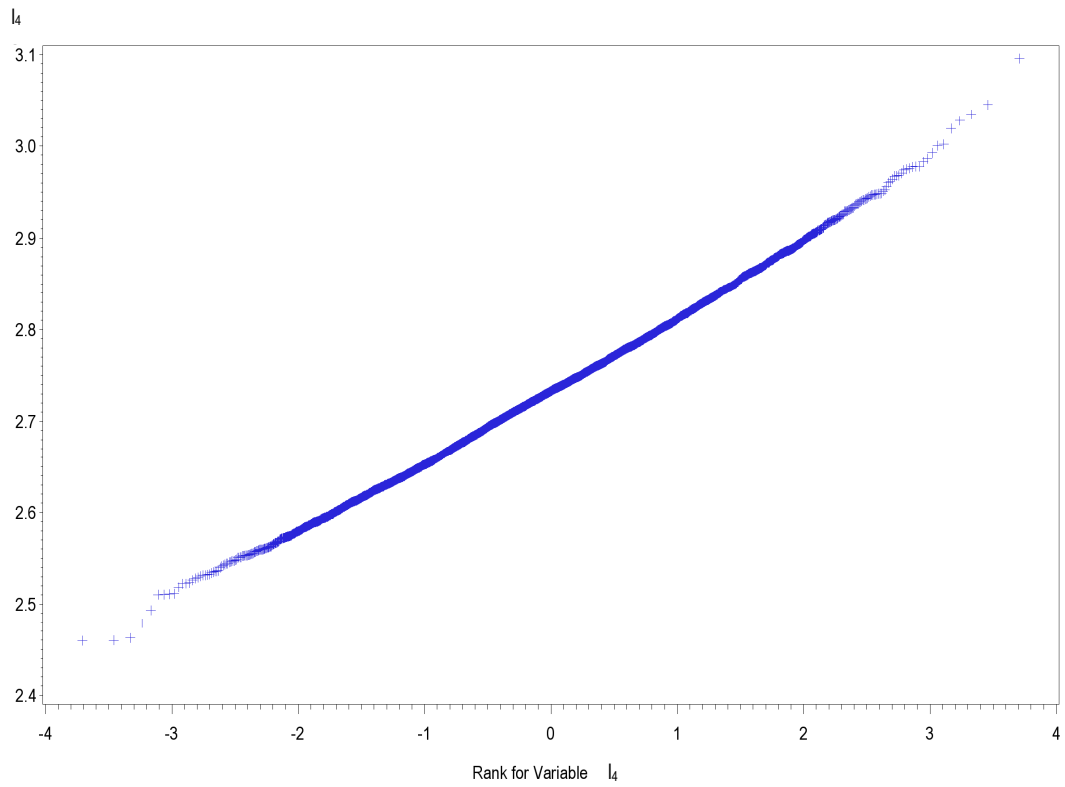
nsub=4800 nvar=60



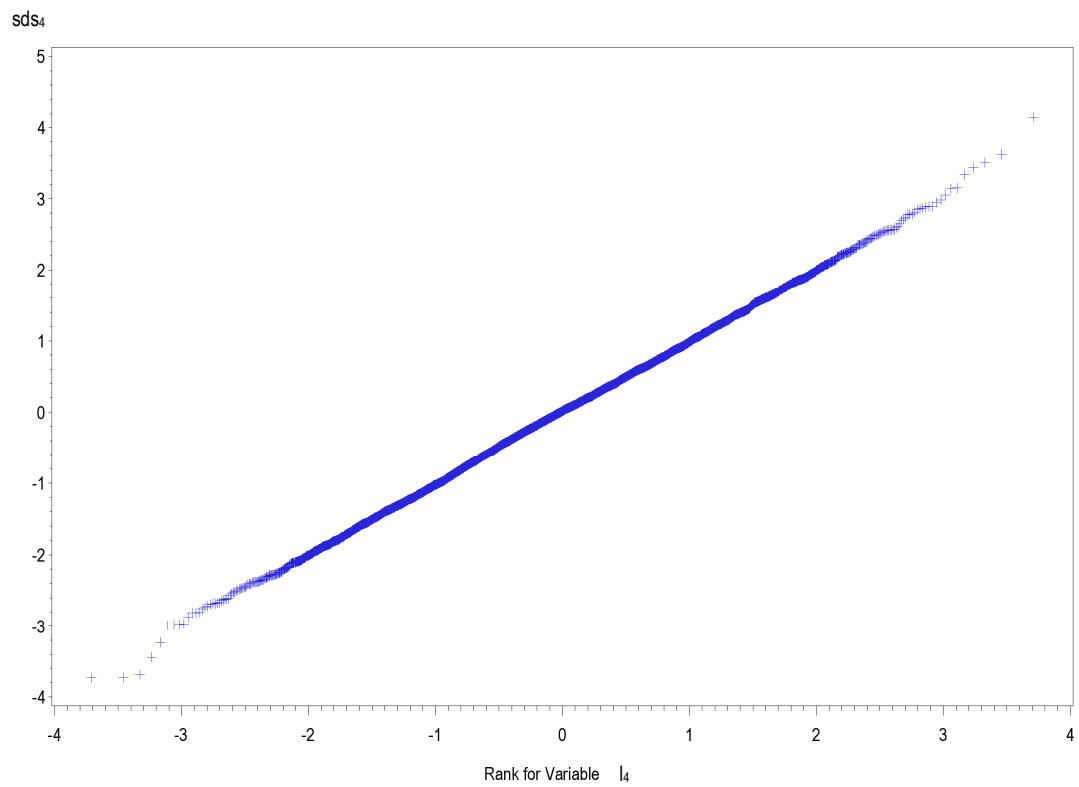
nsub=4800 nvar=60



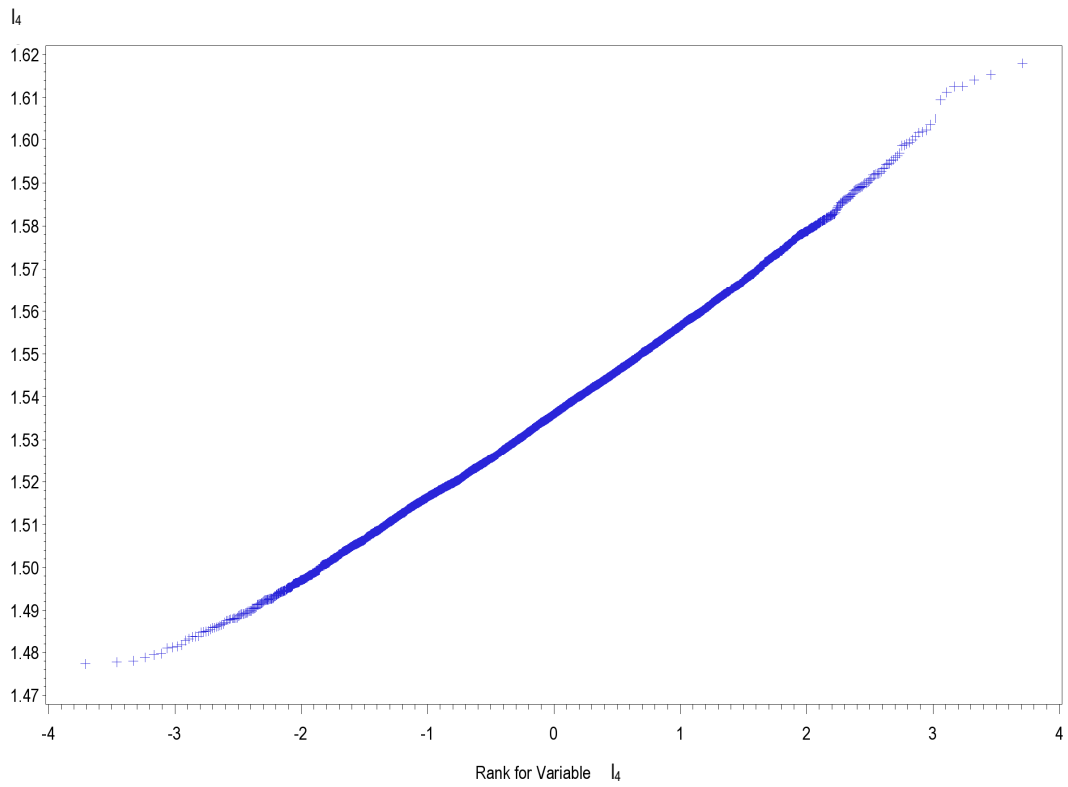
nsub=75 nvar=60



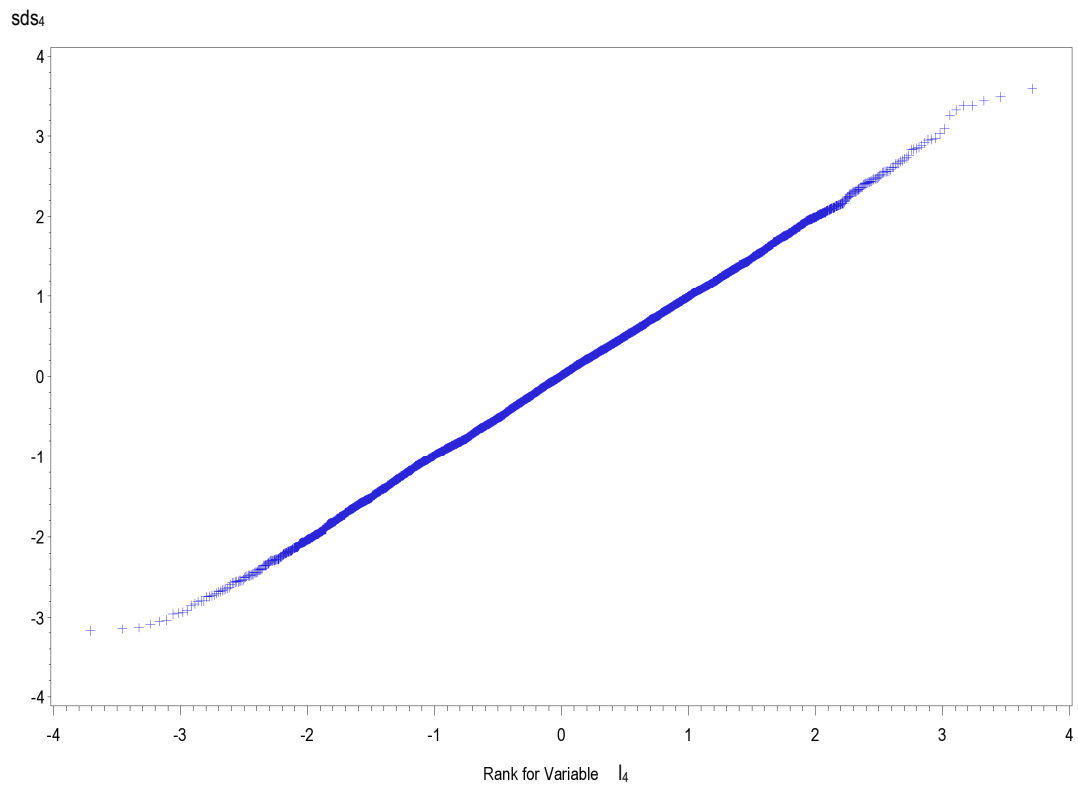
nsub=75 nvar=60



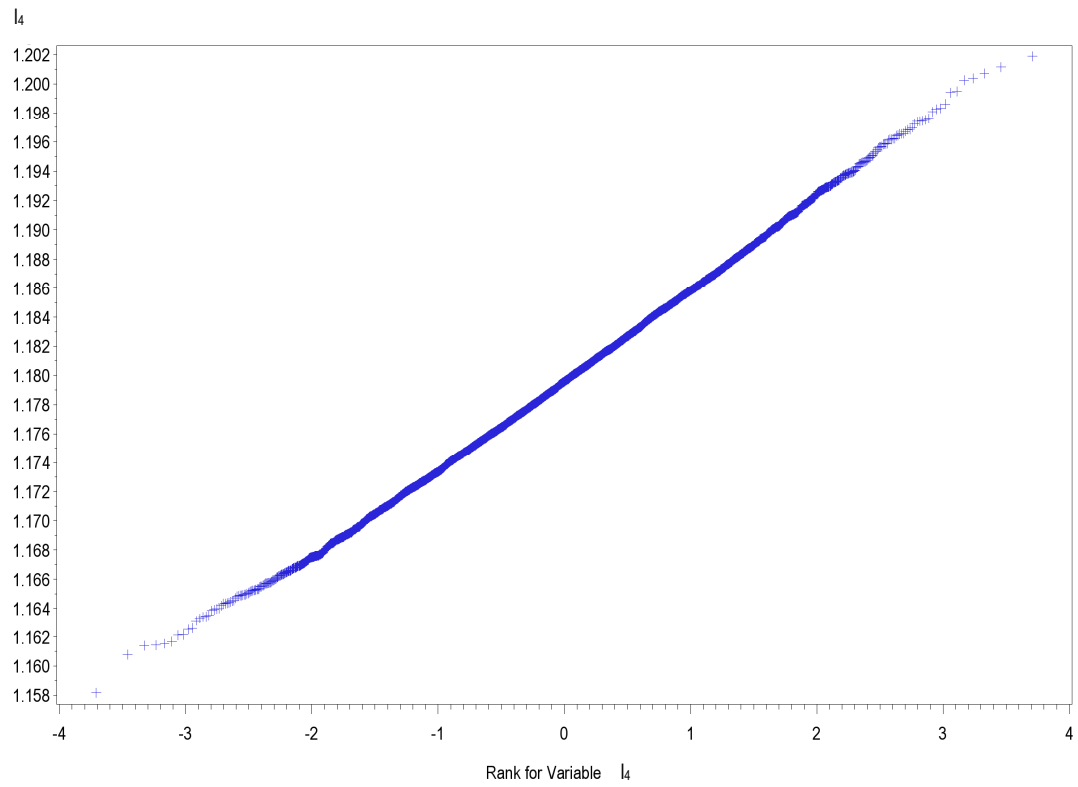
nsub=600 nvar=60



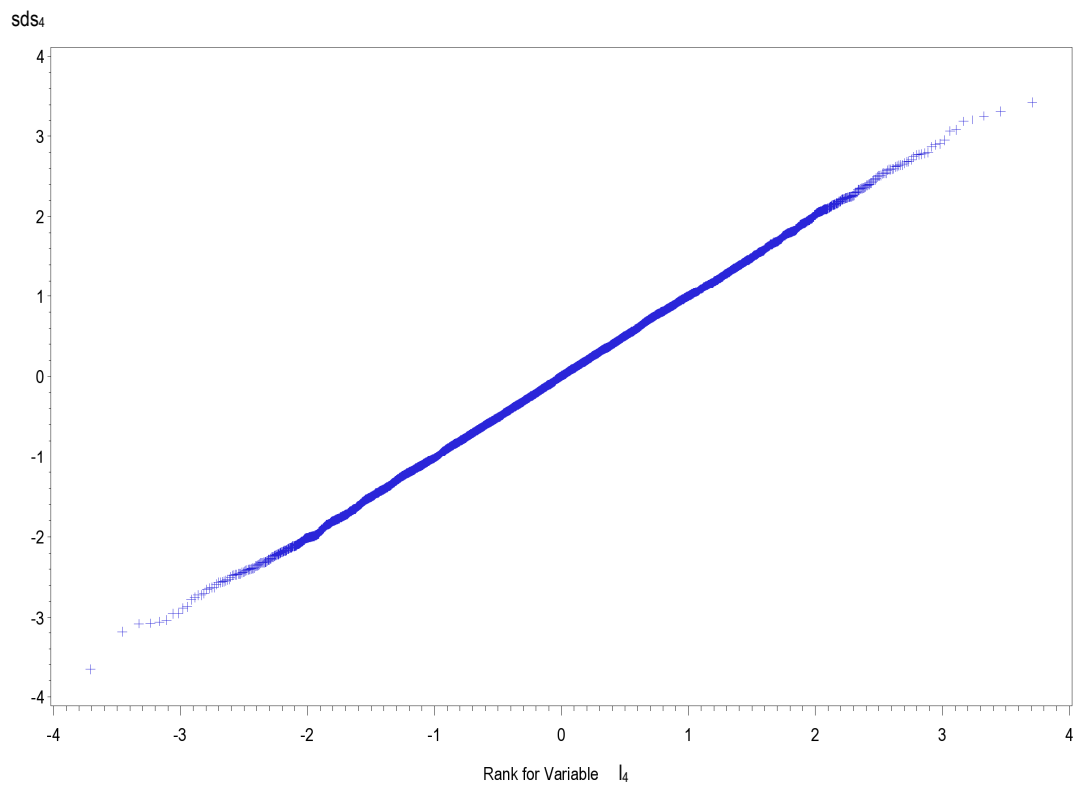
nsub=600 nvar=60



nsub=4800 nvar=60



nsub=4800 nvar=60



6. Bibliografia

1. Bao Z, Pan G, Zhou W. (2011). "Tracy–Widom law for the extreme eigenvalues of sample correlation matrices." Disponibile su arXiv:1110.5208v2 [math.ST].
2. Biondini ME, Mielke PW, Redente EF (1988). "Permutation techniques based on euclidean analysis spaces: a new and powerful statistical method for ecological research." *Coenoses*, 3: 155-174.
3. Boik RT (2003). "Principal component models for correlation matrices." *Biometrika*, 90 (3): 679-701.
4. Boik RT (2004). "Second-order accurate inference on eigenvalues of covariance and correlation matrices.", *Journal of Multivariate Analysis*, 96: 136-171.
5. Boik RT (2013). "Model-based principal components of correlation matrices." *Journal of Multivariate Analysis*, 116: 310-331.
6. Box GEP, Cox DR (1964). "An analysis of transformations." *Journal of the Royal Statistical Society. Series B (Methodological)*: 211-252.
7. Buja A, Eyuboglu N (1992). "Remarks on Parallel Analysis." *Multivariate Behavioral Research*, 27 (4): 509-540.
8. Coste J, Bouée S, Ecosse E, Leplège A, Pouchot J (2005). "Methodological issues in determining the dimensionality of composite health measures using principal component analysis: Case illustration and suggestions for practice." *Quality of Life Research*, 14: 641-654.
9. Cota AA, Longman S, Holden RR, Fekken GC, Xinaris S (1993a). "Interpolating 95th percentile eigenvalues from random data: an empirical example." *Educational and Psychological Measurement*, 53: 585-596.
10. Cota AA, Longman RS, Holden RR, Fekken GC (1993b). "Comparing different methods for implementing Parallel Analysis: a practical index of accuracy." *Educational and Psychological Measurement*, 53: 865-876.
11. Crawford AV, Green SB, Levy R, Lo WJ, Scott L, Svetina D, Thompson MS (2010). "Evaluation of Parallel Analysis methods for determining the number of factors." *Educational and Psychological Measurement*, 70: 885-901.
12. Dinno A (2009a). "Exploring the sensitivity of Horn's Parallel Analysis to the distributional form of random data" *Multivariate Behavioral Research*, 44 (3): 362-388.
13. Dinno A (2009b). "Implementing Horn's Parallel Analysis for principal component analysis and factor analysis." *The STATA Journal*, 9 (2): 291-298.
14. Franklin SB, Gibson DJ, Robertson PA, Pohlmann JT, Fralish JS (1995). "Parallel Analysis: a method for determining significant principal components." *Journal of Vegetation Science*, 6: 99-106.
15. Glorfeld LW (1995). "An improvement on Horn's Parallel Analysis methodology for selecting the correct number of factors to retain." *Educational and Psychological Measurement*, 55: 377-393.

16. Green SB, Levy R, Thompson MS, Lu M, Lo WJ (2012). "A proposed solution to the problem with using completely random data to assess the number of factors with Parallel Analysis." *Educational and Psychological Measurement*, 72: 357-374.
17. Haynes MA, MacGillivray HL, Mengersen KL (1997). "Robustness of ranking and selection rules using generalised g-and-k distributions." *Journal of Statistical Planning and Inference*, 65 (1): 45-66.
18. Hayton JC, Allen DG, Scarpello V (2004). "Factor retention decisions in Exploratory Factor Analysis: a tutorial on Parallel Analysis." *Organizational Research Methods*, 7: 191-205.
19. Jackson DA (1993). "Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches." *Ecology*, 74 (8): 2204-2214.
20. Jiang T (2004). "The limiting distributions of eigenvalues of sample correlation matrices." *Sankhyā: The Indian Journal of Statistics*, 66 (1): 35-48.
21. Johnstone IM (2001). "On the distribution of the largest eigenvalue in principal component analysis." *The Annals of Statistics*, 29 (2): 295-327.
22. Johnstone IM (2006). "High dimensional statistical inference and random matrices." *Proceedings of the International Congress of Mathematicians* (307-333). Madrid, Spain: European Mathematical Society.
23. Jolliffe I (1986). *Principal component analysis*. John Wiley & Sons, Ltd.
24. Kabacoff RI (2000). "Determining the dimensionality of data: a SAS® macro for Parallel Analysis." SUGI 28, Paper 90-28.
25. Keeling KB (2000). "A regression equation for determining the dimensionality of data." *Multivariate Behavioral Research*, 35 (4), 457-468.
26. Keeling KB, Pavur RJ (2004). "A comparison of methods for approximating the mean eigenvalues of a random matrix." *Communications in Statistics – Simulation and Computation*, 33 (4): 945-961.
27. Lautenschlager GJ, Lance CE; Flaherty VL (1989a). "Parallel Analysis criteria: revised equations for estimating the latent roots of random data correlation matrices." *Educational and Psychological Measurement*, 49: 339-345.
28. Lautenschlager GJ (1989b). "A comparison of alternatives to conducting Monte Carlo analyses for determining parallel analysis criteria." *Multivariate Behavioral Research*, 24 (3): 365-395.
29. Ledesma RD, Valero-Mora P (2007). "Determining the number of factors to retain in EFA: an easy-to-use computer program for carrying out Parallel Analysis." *Practical Assessment, Research & Evaluation*, 12 (2): 1-11.
30. Longman RS, Cota AA, Holden RR; Fekken CG (1989). "A regression equation for the Parallel Analysis criterion in Principal Component Analysis: mean and 95th percentile eigenvalues." *Multivariate Behavioral Research*, 24 (1): 59-69.
31. Lorenzo-Seva U, Ferrando PJ (2006). "FACTOR: a computer program to fit the exploratory factor analysis model." *Behavior Research Methods*, 38 (1): 88-91.

32. Matsumoto M, Nishimura T (1998). "Mersenne Twister: A 623-dimensionally equidistributed uniform pseudo-random number generator." *ACM Transactions on Modeling and Computer Simulation*, 8: 3-30.
33. Muirhead RJ (1978). "Latent roots and matrix variates: a review of some asymptotic results." *The Annals of Statistics*, 6 (1): 5-33.
34. Muirhead RJ (1982). *Aspects of multivariate statistical theory*. John Wiley & Sons, Inc.
35. O'Connor BP (2000). "SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test." *Behavior Research Methods, Instruments, & Computers*, 32 (3): 396-402.
36. Patil VH, Singh SN, Mishra S, Donavan DT (2008). "Efficient theory development and factor retention criteria: abandon the 'eigenvalue greater than one' criterion." *Journal of Business Research*, 61: 162-170.
37. Peres-Neto PR, Jackson PA, Somers KM (2005). "How many principal components? stopping rules for determining the number of non-trivial axes revisited." *Computational Statistics & Data Analysis*, 49: 974-997.
38. Pillai NS, Yin J (2012). "Edge universality of correlation matrices." *The Annals of Statistics*, 40 (3): 1737-1763.
39. Ricci C, Milani S (2007). "Ci si può fidare dei risultati dell'analisi delle componenti principali?" *Atti del IV Congresso Nazionale della Società di Statistica Medica ed Epidemiologia Clinica (SISMEC)*: Monreale (Palermo).
40. SAS Institute Inc. 2008. SAS/STAT® 9.2 User's Guide. Cary, NC: SAS Institute Inc.
41. Silverstein AB (1990). "Update on the Parallel Analysis criterion for determining the number of principal components." *Psychological Reports*, 67: 511-514.
42. Thompson B, Daniel LG (1996). "Factor analytic evidence for the construct validity of scores: a historical overview and some guidelines." *Educational and Psychological Measurement*, 56: 197-208.
43. Timmerman ME, Lorenzo-Seva U (2011). "Dimensionality assessment of ordered polytomous items with Parallel Analysis." *Psychological Methods*, 16 (2): 209-220.
44. Turner NE (1998). "The effect of common variance and structure pattern on random data eigenvalues: implications for the accuracy of Parallel Analysis." *Educational and Psychological Measurement*, 58: 541-568.
45. Velicer WF, Eaton CA, Fava JL (2000). "Construct explication through Factor or Component Analysis: a review and evaluation of alternative procedures for determining the number of factors or components." In Goffin RD, Helmes E (Eds.), *Problems and solutions in human assessment: honoring Douglas N. Jackson at seventy*. Boston: Kluwer. (41-71).
46. Watkins MW (2006). "Determining Parallel Analysis criteria." *Journal of Modern Applied Statistical Methods*, 5 (2): 344-346.
47. Zwick WR, Velicer WF (1986). "Comparison of five rules for determining the number of components to retain." *Psychological Bulletin*, 99 (3): 432-442.