UNIVERZA NA PRIMORSKEM
FAKULTETA ZA MATEMATIKO, NARAVOSLOVJE IN
INFORMACIJSKE TEHNOLOGIJE

Zaključna naloga

Uporaba permutacijskih testov za primerjavo povprečij dveh populacij

(Using permutation testing for comparison two population means)

Ime in priimek: Blaž Batagelj
Študijski program: Matematika v ekonomiji in financah
Mentor: doc. dr. Rok Blagus

Koper, avgust 2017

Ključna dokumentacijska informacija

Ime in PRIIMEK: Blaž BATAGELJ

Naslov zaključne naloge: Uporaba permutacijskih testov za primerjavo povprečij dveh populacij

Kraj: Koper

Leto: 2017

Število listov: 74 Število slik: 16

Število prilog: 1 Število strani prilog: 16 Število referenc: 7

Mentor: doc. dr. Rok Blagus

Ključne besede: dvovzorčno testiranje hipotez, normalna porazdelitev, Pareto porazdelitev, permutacijski test, simulacije

Math. Subj. Class. (2010): 62

Izvleček:
Ena izmed najbolj pogosto uporabljenih statističnih metod je testiranje razlike med dvema povprečjema. V ta namen se običajno predpostavi normalno porazdelitev in homogenost varianc, podatke pa se analizira s testom t za dva neodvisna vzorca. Žal finančni podatki pogosto niso porazdeljeni normalno, ampak izkazujejo neko pozitivno asimetrično obliko porazdelitve, kot je na primer Pareto porazdelitev. V zaključni nalogi bomo opisali nekaj osnovnih statističnih definicij. Predstavili bomo test t in test razmerja verjetij, s katerima testiramo zadale hipoteze pri primerjanju dveh povprečij. Dokazali bomo, da sta test t in test razmerja verjetji, v primeru normalno porazdeljenih populacij, ekvivalentna. S pomočjo izrekov bomo izračunali porazdelitve testnih statistik, kar nam bo omogočilo določiti kdaj bo test zavrnil ničelno domnevo. Predstavili bomo Pareto porazdelitev in preverili lastnosti testa t ter izpeljali testno statistiko razmerja verjetij za dve Pareto porazdeljeni populaciji, kjer nas bosta zanimala velikost in moč testa. Tu bomo povedali zakaj so permutacijski testi sploh uporabni ter kako se jih izvede. Ogledali si bomo tudi ali je mogoče dobiti bolj zanesljive rezultate z uporabo permutacijskih testov. Rezultati bodo temeljili na simuliranih podatkih. Sledila bo obrazložitev algoritma, s katerim bomo izvedli simulacije. Algoritem bomo zagnali v programu R. Rezultate bomo grafično in besedno tudi obrazložili.

Key words documentation

Name and SURNAME: Blaž BATAGELJ

Title of final project paper: Using permutation testing for comparison two population means

Place: Koper

Year: 2017

Number of pages: 74 Number of figures: 16

Number of appendices: 1 Number of appendix pages: 16 Number of references: 7

Mentor: Assist. Prof. Rok Blagus, PhD

Keywords: two-sampled testing hypotheses, normal distribution, Pareto distribution, permutation test, simulations

Math. Subj. Class. (2010): 62

Abstract: One of the most used statistical method is testing differences between two means. For this purpose we usually assume normal distribution and homogeneity of variance while data are analysed by t-test for two independent samples. Unfortunately, financial data are not often normally distributed but they show some positive asymmetric distribution form for example Pareto distribution. In this diploma we will describe some basic statistical definitions. We will present t-test and likelihood ratio test which are used for testing hypothesis for testing two means. We will prove that t-test and likelihood ratio test are equivalent in case of normally distributed populations. With the help of theorems we will calculate distributions of test statistics which will enable decide when will test reject the null hypothesis. We will present Pareto distribution, check the properties of t-test and derive test statistic for likelihood ratio for two Pareto distributed populations where we will be interested in size and power of the test. Here we will say why are permutation tests useful and how to perform them. We will see if we can get more reliable results by using permutation tests. The results will be based on simulated data. Following by the explanation of algorithm which we will perform simulations with. The algorithm will be run in program R. The results will be graphically and verbal explained.

Zahvala

Zahvaljujem se svojemu mentorju doc.dr.Roku Blagusu za pomoč ter usmerjanje pri pisanju zaključne naloge. Z mirno in jasno razlago je poskrbel za motivacijo in dobro razumevanje ozadja zaključnega dela.

Kazalo vsebine

Kazalo slik

Kazalo prilog

1 Uvod

Pri statistiki pogosto želimo preveriti resničnost dveh nasprotnih si hipotez oziroma ali sta dve spremenljivki povezani in kako močan vpliv imata ena na drugo. V zaključni nalogi nas bo zanimala primerjava povprečij med dvema populacijama in na ta način si bomo hipotezi tudi zadali. Pred testiranjem hipotez bomo pojasnili statistične definicije za lažjo predstavo kaj sploh žeimo izračunati in kako rezultate interpretirati ter izreke, ki jih bomo pri testiranju uporabili. Predstavili bomo test razmerja verjetij, t-test in permutacijski test, s katerimi se v statistiki testira razlike med dvema populacijama in pojasnili njihove lastnosti. Z omenjnimi test bomo v simulacijah tudi mi testirali naše hipoteze. Pred samim testiranjem bomo morali določiti predpostavke o porazdelitvah obeh populacij. Najprej bomo predpostavljali, da sta populaciji porazdeljeni normalno, ki velja za najbolj pogosto porazdelitev. Nato bomo predpostavljali še, da sta populaciji porazdeljeni Pareto, s katero se pogosto srečujemo v ekonomiji. Lastnosti obeh porazdelitev bomo v teoretičnem delu tudi predstavili. Tu bomo poudarili prednosti in slabosti testa razmerja verjetij ter t-testa pri obeh porazdelitvah in zakaj s permutacijskim testom te slabosti odpravimo. Sledila bo razlaga algoritma in predstavitev programa R, s katerim bomo izvedli simulacije. Rezultate simulacij bomo besedno in grafično interpretirali.

2 Splošno o testiranju hipotez

Ko želimo preveriti ali neka trditev v populaciji velja, moramo zaradi nedostopnosti podatkov celotne populacije izvesti vzorec, na katerem testiramo zadale hipoteze. Statistična hipoteza je trditev o vrednosti enega ali več populacijskih parametrov. Lahko pa je tudi trditev o celotni verjetnostni porazdelitvi. Hipoteza je enostavna, če popolnoma določa porazdelitev. Pri vsakem testiranju hipotez postavimo dve nasprotni si hipotezi. Prvo hipotezo imenujemo ničelna hipoteza in jo označimo s H_0. Drugo hipotezo, ki je nasprotna ničelni, imenujemo alternativna hipoteza in jo označimo s H_A. Na podlagi vzorca želimo izvedeti, ali imamo dovolj razlogov za zavrnitev H_0 in sprejetje H_A. V nasprotnem primeru H_0 ne moremo zavrniti. Pravilu, ki ga uporabimo za sprejetje odločitve ali zavrniti H_0 ali ne, na podlagi vzorca, imenujemo testni postopek. Ker testni postopek izvajamo na vzorcu, lahko pri tem naredimo dve vrsti napak. Prvo napako imenujemo napaka I. vrste. To napako naredimo, če zavrnemo H_0, ko ta v resnici velja. Drugo napako imenujemo napaka II. vrste. Naredimo jo, če ne zavrnemo H_0, ko je ta v resnici napačna. Ti dve napaki nastaneta, ker testiramo hipotezi na podlagi vzorca, torej je skoraj nemogoče, da bi ju popolnoma odpravili. Želimo pa si, da bi bili ti dve napaki čim manjši. Verjetnost, da naredimo napako I vrste, imenujemo tudi stopnja značilnosti in jo označimo z α. Pove nam, kolikokrat bomo na dolgi rok nepravilno zavrnili H_0, ki sicer velja, če testni postopek ponavljamo na različnih vzorcih iz iste populacije. Najpogosteje uporabljene vrednosti za α sta 0.01 in 0.05. Verjetnost napake II vrste označimo z β. Pove nam, koliko krat bomo na dolgi rok nepravilno sprejeli H_0, ko je ta napačna. Verjetnost zavrnitve H_0, ko je ta napačna imenujemo moč testa. Izračunamo jo kot $1 - \beta$. Po vseh končanih izračunih dobimo končno vrednost testiranja hipotez, ki jo imenujemo testna statistika. Glede na njeno vrednost določimo v katero izmed hipotez bomo bolj verjeli. Množici vrednosti testne statistike, pri kateri zavrnemo H_0 pravimo območje zavrnitve, množici vrednosti pri kateri ne zavrnemo H_0 pa območje sprejema. Tu je ključno, da poznamo tudi porazdelitev testne statistike. Pogosto eksaktne porazdelitve testne statistike ni mogoče najti. V tem primeru izračunamo asimptotsko porazdelitev, ki opisuje približno porazdelitev testne statistike pri dovolj velikem vzorcu. Poleg testne statistike poznamo še p-vrednost ali opazovana stopnja značilnosti. Deifnirana je kot verjetnost, ki je izračunana ob predpostavki pravilne H_0, da je dobljena testna statistika vsaj to-

liko ali bolj kontradiktorna H_0, kot dejanska vrednost, ki smo jo dobili. Ker bomo v nadaljevanju H_0 zavračali samo za velike vrednosti testne statistike lahko formalno definicijo p-vrednosti zapišemo kot:

$$\alpha = \sup_{H_0} P_0(T > c),$$

$$p = \sup_{H_0} P_0(T > t),$$

kjer je slučajna spremenjivka T vrednost testne statistike, c in t pa izbrani vrednosti. Velja naslednje:

- H_0 zavrnemo, če $p - vrednost \leq \alpha$
- H_0 ne zavrnemo, če $p - vrednost > \alpha$ [1]

Če je dejanska verjetnost napake I. vrste manjša od željene, potem takemu testu pravimo konservativni test. Če je dejanska verjetnost napake I. vrste večja od željene, potem takemu testu pravimo liberalni test. Pri predstavitvi simulacij bomo najprej preverili velikost testa. Moč testa pa bomo preverjali samo za konservativne teste. Povedali smo že, da pri testiranju hipotez vedno postavimo dve nasprotni si hipotezi H_0 in H_A. Testni statistiki

$$P_0(x)/P_A(x),$$

kjer je $P_0(x)$ verjetnost dogodka x pod veljavno H_0, $P_A(x)$ pa verjetnost dogodka x pod veljavno H_A, pravimo test razmerja verjetij. Glede na Neyman-Pearsonovo lemo, je test razmerja verjetij optimalen za testiranje dveh enostavnih hipotez, kar pomeni, da ima največjo moč med vsemi testi. [3] Test razmerja verjetij bo v prid H_0, ko bo $P_0(x)/P_A(x) > 1$ in v prid H_A, ko bo $P_0(x)/P_A(x) < 1$. Včasih se test razmerja verjetij piše kot $P_A(x)/P_0(x)$. V tem primeru moramo tudi končno vrednost obratno interpretirati. [1]

2.1 Posplošen test razmerja verjetij

V primeru, ko hipotezi nista enostavni, testa razmerja verjetij ne moremo uporabiti. Še več, optimalni test za preverjanje dveh sestavljenih hipotez niti ne obstaja. Lahko pa uporabimo test, ki je kljub temu, da nam noben izrek ne zagotavlja njegove optimalnosti, zelo koristen. Imenujemo ga posplošeni test razmerja verjetij. Za razumevanje le-tega moramo najprej razumeti metodo največjega verjetja, ki jo uporabljamo pri ocenjevanju parametrov. Če so $X_1, \ldots, X_n$ slučajne spremenljivke porazdeljene s skupno gostoto $f(x_1, \ldots, x_n|\theta)$, potem je verjetje za θ, kot funkcija dobljenih vrednosti na

vzorcu $X_i = x_i$ za $i = 1, \ldots, n$, definirano kot $lik(\theta) = f(x_1, \ldots, x_n|\theta)$. Največja verjetnostna cenilka za θ je tista vrednost parametra θ, pri kateri je verjetje največje. Če so slučajne spremenljivke $X_1, \ldots, X_n$ med sabo neodvisne in enako porazdeljene, velja, da je njihova skupna gostota enaka produktu posameznih gostot, torej lahko zapišemo

$$lik(\theta) = \prod_{i=1}^{n} f(X_i|\theta)$$

Maksimirati tako funkcijo je včasih lahko zapleteno, zato ponavadi raje maksimiramo njen naravni logaritem

$$l(\theta) = \sum_{i=1}^{n} \ln[f(X_i|\theta)].$$

Pogosto srečamo hipoteze, ki določajo vrednost parametrov porazdelitve slučajnih spremenljivk $X_1, \ldots, X_n$, glede na dobljene vrednosti $x_1, \ldots, x_n$ na vzorcu. Tako lahko H_0 določa $\theta \in \omega_0$, kjer je ω_0 podmnožiča vseh možnih vrednosti za θ, H_A pa določa $\theta \in \omega_A$, kjer je $\omega_A = \Omega \setminus \omega_0$, če je $\Omega = \omega_0 \cup \omega_A$ in $\omega_0 \cap \omega_A = \emptyset$. Posplošeni test razmerja verjetij definiramo kot razmerje verjetij ocenjeni s tistim parametrom θ, ki ju maksimira

$$\Lambda^* = \frac{max_{\theta \in \omega_0}[lik(\theta)]}{max_{\theta \in \omega_A}[lik(\theta)]}.$$

H_0 zavrnemo za majhne vrednosti Λ^*. Zaradi tehničnih razlogov raje uporabimo testno statistiko

$$\Lambda = \frac{max_{\theta \in \omega_0}[lik(\theta)]}{max_{\theta \in \Omega}[lik(\Omega)]}$$

Upoštevajmo, da $\Lambda = min(\Lambda^*, 1)$, torej bomo za majhne vrednosti Λ^* H_0 zavrnili tudi pri majhnih vrednostih Λ. Območje zavrnitve za test razmerja verjetij je sestavljeno iz majhnih vrednosti Λ, na primer, za vse $\Lambda \leq \lambda_0$, kjer je λ_0 izbran tako, da $P(\Lambda \leq \lambda_0|H_0) = \alpha$, željena stopnja značilnosti. Ker zadali hipotezi pri posplošenem testu razmerja verjetij nista enostavni, o porazdelitvi Λ tudi ne vemo ničesar. Naslednji izrek, ki ga je napisal ameriški matematik Samuel Stanley Wilks, nam pove asimptotsko porazdelitev Λ.

Izrek 2.1. *(Wilksov izrek) Ob dani gostoti in predpostavki, da H_0 velja, porazdelitev $-2ln(\Lambda)$, ko $n \to \infty$, konvergira proti χ^2 s $dim\,\Omega - dim\,\omega_0$ stopinjami prostosti, kjer sta $dim\,\Omega$ in $dim\,\omega_0$ število prostih parametrov pod Ω ter ω_0.*

[3]

Dokaz. Dokaz izreka lahko najdemo v [4]. □

3 Normalna porazdelitev in t-test

3.1 Lastnosti normalne porazdelitve

Pri iskanju porazdelitve slučajnih spremenljivk največkrat naletimo na normalno ali Gaussovo porazdelitev. Uporabna je predvsem zaradi centralnega limitnega izreka, ki nam pod določenimi pogoje pove, da vzorčno povprečje slučajnih spremenljivk, izbrano neodvisno iz neodvisnih porazdelitev konvergira v porazdelitvi k normalni porazdelitvi, ko je število observacij dovolj veliko. Normalna porazdelitev ima zvonasto obliko, zato lahko včasih naletimo na ime zvonasta krivulja. Porazdelitev vsebuje parametra $\mu \in \mathbb{R}$, kateri označuje populacijsko povprečje, ter $\sigma^2 > 0$, kateri označuje populacijsko varianco. Oznaka za normalno porazdelitev je $N(\mu, \sigma^2)$. Njena gostota je

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Porazdelitvena funkcija slučajne spremenljivke porazdeljene normalno je

$$F_X(x) = \frac{1}{2}\left[1 + erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right],$$

kjer je $erf(x)$ funkcija napake. Pričakovana vrednost slučajne spremenljivke porazdeljene normalno je

$$E(X) = \mu.$$

Varianca slučajne spremenljivke porazdeljene normalno je

$$Var(X) = \sigma^2.$$

[5]

3.2 Posplošen test razmerja verjetij pri normalni porazdelitvi

V naslednjem primeru bom predstavil uporabo posplošenega testa razmerij za podatke, kjer predpostavljamo normalno porazdelitev. Naj bodo podatki $X_1, \ldots, X_n$ med sabo

neodvisni in enako porazdeljeni s povprečjem μ in varianco σ^2, kjer je σ poznana. Zadamo si ničelno in alternativno hipotezo:

$$H_0 : \mu = \mu_0$$

$$H_A : \mu \neq \mu_0,$$

kjer je μ_0 nek dana vrednost. Definirajmo množice parametrov $\omega_0 = \mu_0$, $\omega_A = \{\mu | \mu \neq \mu_0\}$ in $\Omega = -\infty < \mu < \infty$. Ker pod ω_0 natančno določamo μ, lahko števec razmerja verjetij zapišemo kot

$$\frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{1}{2\sigma^2}\sum_{i=0}^{n}(X_i-\mu_0)^2}$$

Za imenovalec razmerja verjetij moramo poiskati tisti μ, za katerega bo verjetje največje. Tak μ dobimo z uporabo metode največjega verjetja, za katero je znano, da bomo za rezultat dobili $\mu = \overline{X}$. Torej, bo imenovalec razmerja verjetij

$$\frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\frac{1}{2\sigma^2}\sum_{i=0}^{n}(X_i-\overline{X})^2}$$

Testna statistika razmerja verjetij je torej

$$\Lambda = e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(X_i-\mu_0)^2 - \sum_{i=1}^{n}(X_i-\overline{X})^2\right]}$$

Opazimo, da H_0 zavrnemo za majhne vrednosti Λ. Kar pomeni, da H_0 zavrnemo za velike vrednosti

$$-2\ln\Lambda = \frac{1}{\sigma^2}\left[\sum_{i=1}^{n}(X_i-\mu_0)^2 - \sum_{i=1}^{n}(X_i-\overline{X})^2\right]$$

Za preureditev izraza uporabimo naslednjo enakost:

$$\sum_{i=1}^{n}(X_i-\mu_0)^2 = \sum_{i=1}^{n}(X_i-\overline{X})^2 + n(\overline{X}-\mu_0)^2$$

Dobljen test razmerja verjetij zavrne za velike vrednosti

$$-2\ln\Lambda = \frac{n(\overline{X}-\mu_0)}{\sigma^2}$$

Vemo, da je pod H_0, $\overline{X} \sim N(\mu_0, \frac{\sigma^2}{n})$, zato velja $\frac{\sqrt{n}\overline{X}-\mu_0}{\sigma} \sim N(0,1)$. Prepoznamo, da je porazdelitev testne statistike pod H_0 kvadrirana standardna normalna porazdelitev, za katero velja $Z^2 \sim \chi_1^2$. Torej,

$$-2\ln\Lambda \sim \chi_1^2$$

Ta porazdelitev testne statistike je eksaktna. Če je α izbrana stopnja značilnosti, bo test zavrnil H_0, ko

$$\frac{n(\overline{X}-\mu_0)^2}{\sigma^2} > \chi_1^2(\alpha),$$

kjer je $\chi_1^2(\alpha)$ vrednost, ki jo odčitamo iz tabele za hi-kvadrat porazdelitev. V tem primeru pri odčitavanju upoštevamo, da imamo stopnjo značilnosti α pri 1 stopinji prostosti. [3]

3.3 Testiranje razlik med dvema normalno porazdeljena populacijama - t-test

Pri testiranju hipotez pogosto preverjamo razlike med dvema populacijama. To pogosto storimo tako, da primerjamo razliko njunih povprečij, torej želimo s pomočjo dveh med sabo neodvisnih vzorcev preveriti vrednost $\mu_1 - \mu_2$, kjer sta μ_1 povprečna vrednost v prvi populaciji in μ_2 povprečna vrednost v drugi populaciji. Torej bomo iskane vrednosti dobili za vsak vzorec posebej. Tu predpostavimo, da je prvi vzorec $X_1, \ldots, X_n$ izbran iz populacije porazdeljene normalno s povprečjem μ_X, drugi vzorec $Y_1, \ldots, Y_m$ pa je izbran iz neke druge populacije porazdeljene normalno s povprečjem μ_Y. Obe populaciji imata enako varianco σ^2. Po metodi največjega verjetja je cenilka za $\mu_X - \mu_Y$ enaka $\overline{X} - \overline{Y}$. To lahko izrazimo kot linearno kombinacijo neodvisnih normalno porazdeljenih slučajnih spremenjivk. Iz tega sledi, da je tudi slučajna spremenjivka $\overline{X} - \overline{Y}$ porazdeljena normalno:

$$\overline{X} - \overline{Y} \sim N\left[\mu_X - \mu_Y, \sigma(\frac{1}{n} + \frac{q}{m})\right]$$

Izrek 3.1. *Če predpostavimo, da je H_0 pravilna in obe opazovani populaciji porazdeljeni normalno in z enako varianco, potem je standardizirana spremenljivka*

$$t = \frac{\overline{x}_1 - \overline{x}_2 - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

eksaktno porazdeljena s t porazdelitvijo z $m + n - 2$ stopinjami prostoti. Takemu testu pravimo t-test.

Dokaz. Dokaz izreka lahko najdemo v [3]. □

Testna statistika, na podlagi katere se bomo odločili ali zavrniti $H_0 : \mu_x = \mu_y$ je

$$t = \frac{\overline{X} - \overline{Y}}{s\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

Porazdelitev te testne statistike pod H_0 je t porazdelitev z $m + n - 2$ stopinjami prostoti. [3]

3.3.1 Ekvivalenca t-testa in posplošenega testa razmerja verjetij

V naslednjem postopku bomo pokazali, da sta t-test in posplošen test razmerja verjetij ekvivalentna. Definirajmo množico Ω, ki predstavlja vse možne vrednosti vseh parametrov:

$$\Omega = \{-\infty < \mu_x < \infty, -\infty < \mu_y < \infty, 0 < \sigma < \infty\}$$

Ti parametri nam niso poznani, zato definiramo vektor neznanih parametrov kot $\theta = (\mu_x, \mu_y, \sigma)$. Sedaj lahko postavimo ničelno hipotezo:

$$H_0 : \theta \in \omega_0, \quad \omega_0 = \{\mu_x = \mu_y, 0 < \sigma < \infty\}$$

. Verjetje dveh vzorcev $X_1, \ldots, X_n$ in $Y_1, \ldots, Y_m$ zapišemo kot

$$lik(\mu_x, \mu_y, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(X_i-\mu_x)^2}{2\sigma^2}} \prod_{j=1}^{m} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(Y_j-\mu_y)^2}{2\sigma^2}}$$

Izraz logaritmiramo in dobimo

$$l(\mu_x, \mu_y, \sigma^2) = -\frac{(m+n)}{2} \ln 2\pi - \frac{(m+n)}{2} \ln 2\sigma^2 - \frac{1}{2\sigma^2} \left[\sum_{i=1}^{n} (X_i - \mu_x)^2 + \sum_{j=1}^{m} (Y_i - \mu_y)^2 \right]$$

Poiskati moramo maksimum tega verjetja pod ω_0 in Ω ter izračunati razmerje obeh maksimiranih verjetij. Ker pod ω_0 trdimo, da sta povprečji obeh vzorcev enaki $\mu_x = \mu_y$, imamo tu dva neznana parametra μ_0 in σ_0^2, ki ju moramo oceniti. Torej, naši podatki so normalno porazdeljeni s povprečjem μ_0 in varianco σ_0^2 Velikost vzorca je $m+n$. Cenilke za μ_0 in σ_0^2 dobimo z uporabo metode največjega verjetja in s tem maksimiramo verjetje:

$$\hat{\mu}_0 = \frac{1}{m+n} \left(\sum_{i=1}^{n} X_i + \sum_{j=1}^{m} Y_j \right)$$

$$\hat{\sigma}_0^2 = \frac{1}{m+n} \left[\sum_{i=1}^{n} (X_i - \hat{\mu}_0)^2 + \sum_{j=1}^{m} (Y_j - \hat{\mu}_0)^2 \right]$$

Po nekaj korakih dobimo vrednost logaritmiranega verjetja:

$$l(\hat{\mu}_0, \hat{\sigma}_0^2) = -\frac{m+n}{2} \ln(2\pi) - \frac{m+n}{2} \ln(\hat{\sigma}_0^2) - \frac{m+n}{2}$$

Poiskati moramo še maksimum verjetja pod Ω. V nasprotju z ω_0, tu trdimo, da sta povprečji obeh vzorcev različni $\mu_x \neq \mu_y$. Torej, imamo tri neznane parametre μ_x, μ_y in σ_0^2, ki jih moramo oceniti:

$$\sum_{i=1}^{n} (X_i - \hat{\mu}_x) = 0 \quad \Rightarrow \quad \hat{\mu}_x = \overline{X}$$

$$\sum_{j=1}^{m} (Y_j - \hat{\mu}_y) = 0 \quad \Rightarrow \quad \hat{\mu}_y = \overline{Y}$$

$$-\frac{m+n}{2\hat{\sigma}_A^2} + \frac{1}{2\hat{\sigma}_A^4} \left[\sum_{i=1}^{n} (X_i - \hat{\mu}_x)^2 + \sum_{j=1}^{m} (Y_j - \hat{\mu}_y)^2 \right]$$

$$\Rightarrow \qquad \hat{\sigma}_A^2 = \frac{1}{m+n} \left[\sum_{i=1}^{n} (X_i - \hat{\mu}_x)^2 + \sum_{j=1}^{m} (Y_j - \hat{\mu}_y)^2 \right]$$

Ko vstavimo cenilke v logaritmirano verjetje dobimo

$$l(\hat{\mu}_x, \hat{\mu}_y, \hat{\sigma}_A^2) = -\frac{m+n}{2}\ln(2\pi) - \frac{m+n}{2}\ln(\hat{\sigma}_A^2) - \frac{m+n}{2}$$

Sedaj lahko sestavimo razmerje verjetij:

$$\frac{m+n}{2}\ln\left(\frac{\hat{\sigma}_A^2}{\hat{\sigma}_0^2}\right)$$

Opazimo, da bo test razmerja verjetij zavrnil H_0 pri velikih vrednostih

$$\frac{\hat{\sigma}_0^2}{\hat{\sigma}_A^2} = \frac{\sum_{i=1}^{n}(X_i - \hat{\mu}_0)^2 + \sum_{\mathrm{j}=1}^{m}(Y_j - \hat{\mu}_0)^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2 + \sum_{j=1}^{m}(Y_j - \overline{Y})^2}$$

Števec tega razmerja lahko zapišemo na naslednji način:

$$\sum_{i=1}^{n}(X_i - \hat{\mu}_0)^2 = \sum_{i=1}^{m}(X_i - \overline{X})^2 + n(\overline{X} - \hat{\mu}_0)^2$$

$$\sum_{\mathrm{j}=1}^{\mathrm{n}}(Y_j - \hat{\mu}_0)^2 = \sum_{\mathrm{j}=1}^{\mathrm{m}}(Y_j - \overline{Y})^2 + m(\overline{Y} - \hat{\mu}_0)^2$$

Vemo, da

$$\hat{\mu}_0 = \frac{1}{m+n}(n\overline{X} + m\overline{Y}) = \frac{n}{m+n}\overline{X} + \frac{m}{m+n}\overline{Y}$$

Iz tega sledi

$$\overline{X} - \mu_0 = \frac{m(\overline{X} - \overline{Y})}{m+n}$$

$$\overline{Y} - \mu_0 = \frac{n(\overline{Y} - \overline{X})}{m+n}$$

Torej lahko števec verjetja zapišemo kot

$$\sum_{i=1}^{n}(X_i - \overline{X})^2 + \sum_{j=1}^{m}(Y_j - \overline{Y})^2 + \frac{mn}{m+n}(\overline{X} - \overline{Y})^2$$

Testna statistika zavrne pri velikih vrednostih

$$1 + \frac{mn}{m+n}\left(\frac{(\overline{X} - \overline{Y})^2}{\sum_{i=1}^{n}(X_i - \overline{X})^2 + \sum_{j=1}^{m}(Y_j - \overline{Y})^2}\right)$$

oziroma pri velikih vrednostih

$$\frac{|\overline{X} - \overline{Y}|}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2 + \sum_{j=1}^{m}(Y_j - \overline{Y})^2}}$$

Če zanemarimo konstante, ki niso odvisne od podatkov opazimo, da sta dobljeni testni statistiki pri razmerju verjetja in t testu ekvivalentni, kar smo tudi želeli pokazati. [3]

4 Pareto porazdelitev in permutacijski test

4.1 Lastnosti Pareto porazdelitve

Do sedaj smo predpostavljali, da so slučajne spremenljivke porazdeljene normalno. Porazdelitev slučajnih spremenljivk je seveda več in v praksi mnogokrat spremenljivke niso porazdeljene normalno. Od tu naprej bomo namesto normalne porazdelitve predpostavljali Paretovo porazdelitev, katero se pogosto uporablja v ekonomiji. Imenuje so po italjanskemu ekonomistu Vilfredu Paretu. Porazdelitev vsebuje parametra $x_m \in (0, \infty)$, ki je minimalna vrednost, ki jo slučajna spremenljivka X lahko zavzame, ter $\alpha \in (0, \infty)$. Oznaka za Pareto porazdelitev je $Pareto(\alpha, x_m)$. Njena gostota je

$$f_X(x) = \begin{cases} \frac{\alpha x_m^\alpha}{x_i^{\alpha+1}} & x > x_m, \\ 0 & x < x_m \end{cases}$$

Porazdelitvena funkcija slučajne spremenljivke porazdeljene Pareto je

$$F_X(x) = \begin{cases} 1 - (\frac{x_m}{x})^\alpha & x > x_m, \\ 0 & x < x_m \end{cases}$$

Pričakovana vrednost slučajne spremenljivke porazdeljene Pareto je

$$E(X) = \begin{cases} \infty & \alpha \leq 1, \\ \frac{\alpha x_m^\alpha}{\alpha - 1} & \alpha > 1 \end{cases}$$

Varianca slučajne spremenljivke porazdeljene Pareto je

$$Var(X) = \begin{cases} \infty & \alpha \in (1, 2], \\ (\frac{x_m^\alpha}{\alpha-1})^2 \frac{\alpha}{\alpha-2} & \alpha > 2 \end{cases}$$

Če $\alpha \leq 1$, potem pričakovana vrednost in varianca ne obstaja. Poznamo več tipov Pareto porazdelitve. Mi se bomo ukvarjali zgolj s Pareto porazdelitvijo tipa I. [7]

4.2 Posplošen test razmerja verjetij za Pareto porazdelitev

V naslednjem postopku bomo izpeljali testno statistiko razmerja verjetij za dve Pareto porazdeljeni populaciji. Pri testu razmerja verjetij v primeru normalno porazdeljene populacije smo preverjali enakost povprečij v obeh populacijah. V primeru Pareto porazdeljenih populacij pa preverjamo enakost porazdelitve v obeh populacijah. Formalno, pod H_0 postavimo trditev, da sta porazdelitvi obeh populacij enaki, pod H_A pa trdimo nasprotno.

$$H_0 : X \sim Y$$

$$H_A : X \nsim Y$$

Definirajmo množice parametrov $\omega_0 = \{\alpha_0, x_{m_0}\}$ in $\omega_A = \{\alpha_X, \alpha_Y, x_m, y_m\}$. Verjetje dveh vzorcev $X_1, \ldots, X_n$ in $Y_1, \ldots, Y_m$, ki smo ju izbrali iz populacije porazdeljene Pareto zapišemo kot

$$lik(\alpha_X, \alpha_Y, x_m, y_m) = \prod_{i=1}^{n} \frac{\alpha_X x_m^{\alpha}}{x_i^{\alpha_X+1}} \prod_{j=1}^{m} \frac{\alpha_Y y_m^{\alpha}}{x_j^{\alpha_Y+1}}$$

Verjetje logaritmiramo

$$\begin{aligned} l(\alpha_X, \alpha_Y, x_m, y_m) &= n \ln \alpha_X + n\alpha_X \ln x_m - (\alpha_X + 1) \sum_{i=1}^{n} \ln X_i + m \ln \alpha_Y \\ &\quad + m\alpha_Y \ln y_m - (\alpha_Y + 1) \sum_{j=1}^{m} \ln Y_j \end{aligned}$$

V števcu razmerja upoštevamo, da sta porazdelitvi obeh populacij enaki, zato velja $\alpha_X = \alpha_Y$ in $x_m = y_m$. Torej, logaritmirano verjetje pod ω_0 je

$$l(\alpha_0, x_{m_0}) = (n+m) \ln \alpha_0 + (n+m)\alpha_0 \ln x_{m_0} - (\alpha_0 + 1) \left(\sum_{i=1}^{n} \ln x_i + \sum_{j=1}^{m} \ln y_j \right)$$

Tu imamo dva parametra(α_0 in x_{m_0}), ki ju ocenimo po metodi največjega verjetja

$$\hat{\alpha}_0 = \frac{m+n}{\sum_{i=1}^{n} \ln x_i + \sum_{j=1}^{m} \ln y_j - 2(m+n) \ln \hat{x}_{m_0}}$$

$$\hat{x}_{m_0} = \min_{i,j} (x_i, y_j)$$

V imenovalcu razmerja upoštevamo, da sta porazdelitvi obeh populacij različni, kar nam da naslednje logaritmirano verjetje pod ω_A

$$\begin{aligned} l(\alpha_X, \alpha_Y, x_m, y_m) &= n \ln \alpha_X + n\alpha_X \ln x_m - (\alpha_X + 1) \sum_{i=1}^{n} \ln x_i + m \ln \alpha_Y \\ &\quad + m\alpha_Y \ln y_m - (\alpha_Y + 1) \sum_{j=1}^{m} \ln Y_j \end{aligned}$$

Tu imamo štiri parametre(α_X,α_Y,x_m,y_m), ki jih pravtako ocenimo po metodi največjega verjetja

$$\hat{\alpha}_X = \frac{n}{\sum_{i=1}^{n} \ln x_i - n \ln \hat{x}_m}$$

$$\hat{\alpha}_Y = \frac{m}{\sum_{j=1}^{m} \ln y_j - \ln \hat{y}_m}$$

$$\hat{x}_m = \min_i x_i$$

$$\hat{y}_m = \min_j y_j$$

$$\Lambda = \frac{(n+m)\ln\hat{\alpha}_0 + (n+m)\hat{\alpha}_0 \ln \hat{x}_{m_0} - (\hat{\alpha}_0 + 1)\left(\sum_{i=1}^{n} \ln x_i + \sum_{j=1}^{m} \ln y_j\right)}{n\ln\hat{\alpha}_X + n\hat{\alpha}_X \ln \hat{x}_m - (\hat{\alpha}_X + 1)\sum_{i=1}^{n} \ln x_i + m\ln\hat{\alpha}_Y + m\hat{\alpha}_Y \ln \hat{y}_m - (\hat{\alpha}_Y + 1)\sum_{j=1}^{m} \ln Y_j}$$

Test bo H_0 zavračal pri majhnih vrednostih Λ. Posledično bo test H_0 zavrnil pri velikih vrednostih $-2\ln\Lambda$.

$$\begin{aligned} -2\ln\Lambda \;\approx\; -2\Bigg[& (n+m)(\ln\hat{\alpha}_0 + \hat{\alpha}_0 \ln \hat{x}_{m_0}) - \hat{\alpha}_0\left(\sum_{i=1}^{n} \ln x_i + \sum_{j=1}^{m} \ln y_j\right) \\ & -n(\ln\hat{\alpha}_X - \hat{\alpha}_X \ln \hat{x}_{m_A}) + \hat{\alpha}_X \sum_{i=1}^{n} \ln x_i \\ & -m(\ln\hat{\alpha}_Y - \hat{\alpha_Y} \ln \hat{y}_{m_A}) + \hat{\alpha}_Y \sum_{j=1}^{m} \ln y_j \Bigg] \end{aligned}$$

Porazdelitev testne statistike dobimo z uporabo Wilksovega izreka. Pod ω_0 imamo 2 prosta parametra, pod ω_A pa 4 proste parametre, kar pomeni, da imamo 2 stopinji prostosti.

$$-2\ln\Lambda \sim \chi^2_2$$

Ta porazdelitev testne statistike je zgolj asimptotska. Da bi izračunali eksaktno porazdelitev testne statistike, moramo uporabiti drugačno metodo testiranja hipotez. Ena od le-teh je permutacijska metoda ali randomizacija, ki jo bomo predstavili v nadaljevanju. [7]

4.3 Permutacijski test

Permutacijska metoda ali randomizacija je zelo splošen pristop za testiranje statističnih hipotez. Njena porazdelitev je generirana iz podatkov samih. Zagotavlja nam

učinkovito testiranje hipotez, ko podatki niso skladni s predpostavko o njihovi porazdelitvi. Problema neodvisnosti med opazovanji permutacijska metoda ne odpravi, lahko pa jo uporabimo pri opazovanjih, ki niso bila izbrana naključno. Slabost metode je velika računska zahtevnost za preračunavanje velikega števila permutacij in testnih statistik. Z vse hitrejšimi računalniki na trgu se ta slabost odpravlja.

Za ničelno porazdelitev velja, da so ob pravilni H_0 vsi možni pari dveh spremenljivk enako verjetni, da se pojavijo. Pod H_0 vrednosti vektorja x_1 naključno razporedimo, medtem, ko pozicije vrednosti vekotrja x_2 fiksiramo. S tem dosežemo, da ima vsaka vrednost vektorja x_1 enako verjetnost, da bo z določeno vrednostjo x_2 v paru. Ko so pari določeni, izračunamo vrednost testne statistike. Ta postopek ponovimo večkrat, kar pomeni, da bomo na koncu dobili vektor vrednosti testnih statistik. Te vrednosti primerjamo s testno statistiko, ki smo jo izračunali na podlagi nepermutiranih vrednosti. Tako pridemo do cenilke za porazdelitev testne statistike pod H_0. Tako kot pri drugih statističnih testih se tudi pri permutacijskem testu o zavrnitvi H_0 odločimo tako, da primerjamo dobljeno vrednost testne statistike in porazdelitev pod H_0. Če je vrednost testne statistike vprid ničelne hipoteze, torej, da vektorja x_1 in x_2 nista povezana, potem H_0 ne moremo zavrniti. V nasprotnem primeru, če je vrednost testne statistike vprid alternativne hipoteze, H_0 zavrnemo pri določeni stopnji značilnosti ter sprejmemo H_A.

S permutacijskimi testi preverjamo veljavnost porazdelitve, ki je bila pridobljena tako, da smo dane podatke permutirali. To pomeni, da testna statistika nima nikakršne povezave z opazovano populacijo. Zaradi tega ni potrebno, da so podatki izbrani naključno.

Za majhne nabore podatkov lahko izračunamo vse možne permutacije. S tem bi pridobili popolno permutacijsko porazdelitev testne statistike oziroma izvedli ekzakten ali popolni permutacijski test. Za velike nabore podatkov izvedemo vzorčni permutacijski test, zaradi prevelikega števila permutacij. Zelo pomembno je, da se vse možne permutacije lahko kreirajo z enako verjetnostjo.

Pri uporabi permutacijskega testa za primerjanje razlik povprečij med dvema skupinama naključno permutiramo dva objekta obeh skupin namesto dveh spremenljivk. Način permutiranja je odvisen od zadale ničelne hipoteze. Nekateri testi so le preoblikovani od nekega drugega testa. Na primer, primerjava povprečij dveh populacij s t-testom je ekvivalentna primerjavi korelacij med vektorjem opazovanih vrednosti in vektorjem, ki tem vrednostim dodeli populacijo. Ne glede na to katero metodo uporabimo, dobimo isto vrednost testne statistike. Enostavni statistični testi kot so korelacijski koeficient ali razlike povprečij med dvema populacijama so lahko izvedeni s permutiranjem začetnih podatkov. Problem nastane, če imamo v modelu kompleksne povezave med spremenjivkami. V tem primeru smo lahko primorani permutirati

ostanke modela namesto začetnih podatkov. Takemu načinu pravimo permutacije bazirane na modelu. Mi bomo naše podatke testirali tako, da bomo permutirli vektor, ki predstavlja populacijo.

Če pri vzorčnem permutacijskem testu dodamo referenčno vrednost testne statistike k porazdelitvi, prisilimo test, da proizvede ekstremno vrednost. Ta način računanja verjetnosti je pristranski, vendar velja, da je statistično pravilen. Natančnost te verjetnostne cenilke je obrat števila permutacij. Na primer, če smo izračunali 1000 permutacij, bo natančnost verjetnosti 0.001. Torej nam permutacijski test zagotavlja, da p-vrednost ne bo nikoli enaka 0.

Ko se odločamo o številu permutacij, želimo poiskati ravnotežje med natančnostjo cenilke in zmoglivostjo računalnika. Ker so cenilke izračunane na podlagi vzorca in so s tem prisotne napake, čim več permutacij izvedemo, tem bolje je. Za prvo testiranje hipoteze velja, da naj bi 500 do 1000 permutacij bilo zadostno. Za rezultate, ki jih želimo objaviti v javnosti, naj bi izvedli vsaj 10000 permutacij. [2]

5 Simulacije

5.1 Opis algoritma

Za izvedbo simulacij testiranja hipotez potrebujemo ustrezen algoritem. Definirajmo funkcijo, kateri podamo dva vhodna podatka. Prvi vhodni podatek predstavlja izžrebane vrednosti obeh skupin, drugi vhodni podatek pa predstavlja skupino, iz katere je določena vrednost bila izžrebana. Funkcija izvede posplošeni test razmerja verjetij ob predpostavki, da sta opazovani populaciji porazdeljeni Pareto. Za izhodni podatek dobimo vrednost testne statistike. Na podlagi dobljene testne statistike izračunamo p-vrednosti. Upoštevamo, da je testna statistika porazdeljena po hi-kvadrat porazdelitvi (za kar uporabljamo različne stopinje prostosti). Pri permutacijskem testu pa želimo za vsako permutacijo izračunati vektor testnih statistik. To naredimo s pomočjo `for` zanke, v kateri za vsako permutacijo generira naključen vrsti red komponent v vektorju y in izvede posplošen test razmerja verjetij glede na naključni vektor y. Torej, za vsako permutacijo izračunamo drugačno testno statistiko. Ta postopek se ponavlja, dokler ne doseže željenega števila permutacij. Končni rezultat je vektor testnih statistik. P-vrednost je potem določena, kot število permutiranih testnih statistik, ki so večje, ali enake, od testne statistike dobljene na osnovnih podatkih. Sedaj lahko izvedemo simulacije. Definirajmo funkcijo, kateri podamo velikost vzorca in vrednosti paramterov obeh populacij. Poleg tega podamo še število željenih permutacij. Definirati moramo še dva vektorja. V prvi vektor podamo vrednosti spremenljivke, medtem, ko v drugi vektor generiramo število 1 in 2 glede na velikost vzorca prve in druge populacije s čimer določimo pripadnost skupini. Glede na dobljena vektorja izvedemo t test in posplošen test razmerja verjetij za dva populacij porazdeljeni Pareto. Dobljene vrednosti shranimo v nov vektor. Izvedemo tri različne simulacije, pri katerih spreminjamo velikosti vzorca, vrednosti parametrov in razlike parametrov med obema populacijama. Razlike parametrov bomo označevali z Δ. Prva simulacija predstavlja velikost testa pri istih vrednostih parametrov x_m in α za obe populaciji, zato tu ni razlik med parametri obeh populacij. Druga simulacija predstavlja moč testa, kjer se parameter α med obema populacijama razlikuje za Δ. Tretja simulacija predstavlja moč testa, kjer se parameter x_m med obema populacijama razlikuje za Δ.

5.2 Predstavitev programa R

V računalniškem programu R bomo uporabili omenjeni algoritem. R je programski jezik za statistično računanje in grafiko. Razvil ga je John Chambers s sodelavci na ameriški firmi Bell Laboratories. R nam zagotavlja širok izbor statističnih (linearno in nelinearno modeliranje, testiranje hipotez, analiza časovnih vrst, razvrščanje, grozdenje) in grafičnih tehnik. Ena od prednosti programa R je enostavno in kakovostno risanje grafov, vključno z matematičnimi simboli in formulami. S pomočjo programa R lahko manipuliramo in izračunavamo s podatki ter jih predstavimo z grafičnimi prikazi. Vključuje

- učinkovito obdelavo in shranjevanje podatkov,
- zbirko operaterjev za računanje s polji in matrikami,
- velika in skladna zbirka vmesnih orodij za analizo podatkov,
- grafične zmogljivosti za analizo podatkov in prikaz na zaslonu ali na papirju,
- dobro razvit, preprost in učinkovit programski jezik, ki vključuje pogojne izjave, zanke, rekurzivne funkcije ter vhodne in izhodne zmogljivosti

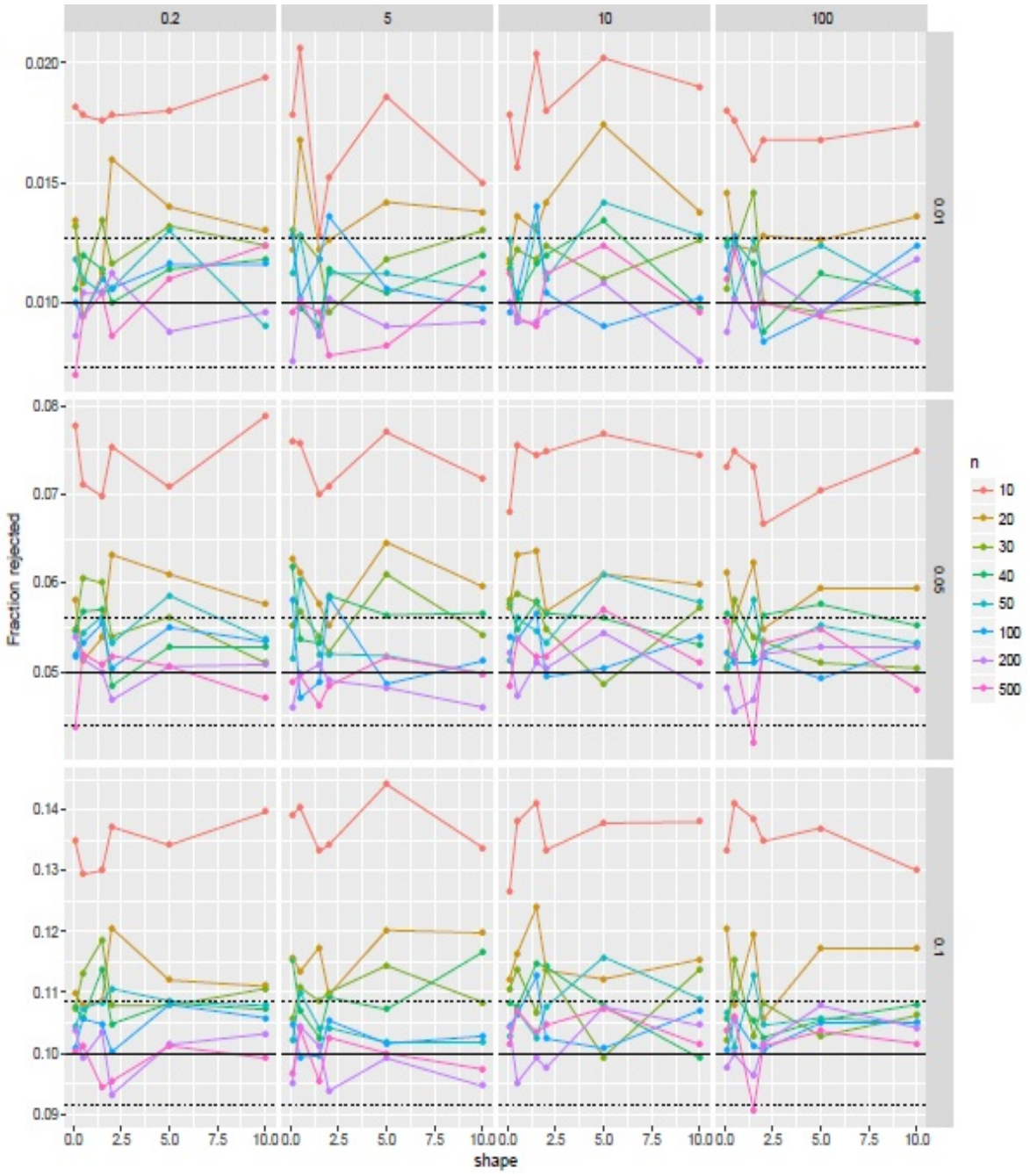
R omogča uporabniku definiranje novih funkcij, enostavno sledenje odločitvam algoritma in povezavo do drugih programskih kod v času izvajanja. [6]

6 Predstavitev rezultatov simulacij

6.1 Velikost testov

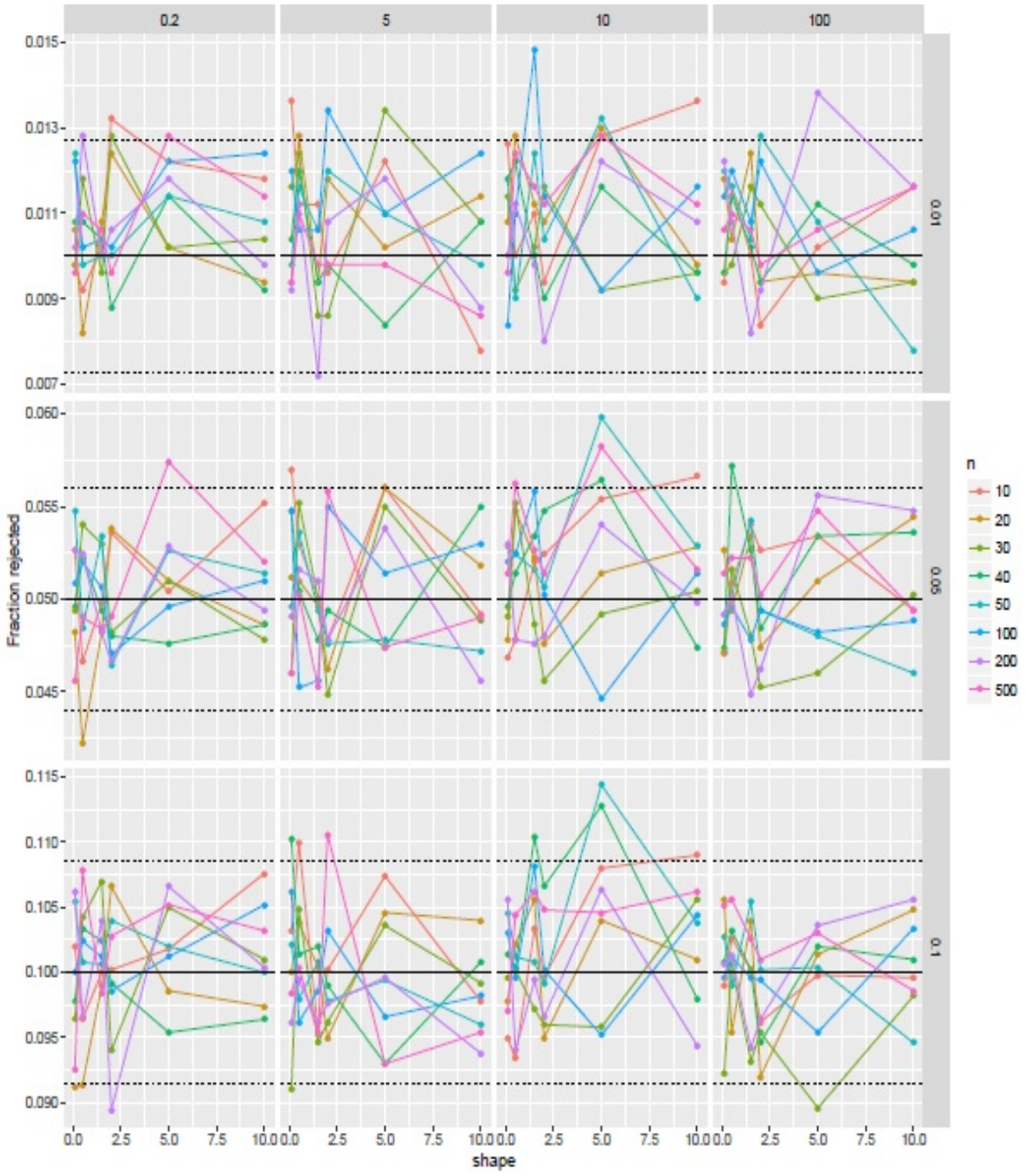
Slika 1 prikazuje velikost testa za dve populaciji porazdeljeni Pareto, kjer smo za izračun p-vrednosti uporabili dve stopinje prostosti. Na x-osi so različne vrednosti parametra α(shape). Na y-osi je delež zavrnitve testa. Različne barve krivulj prikazujejo različne velikosti vzorca. Stolpci grafov se razlikujejo po parametru x_m, vrstice grafov pa po željeni stopnji značilnosti. Zaradi vzorčne napake toleriramo manjša odstopanja od željene stopnje značilnosti, zato sprejemo vse točke grafov, ki so znotraj črtkanih vodoravnih črt. Iz slike 1 razberemo, da se vse točke nahajajo izven črtkanih vodoravnih črt, kar pomeni, da za vsako točko velja, da je dejanska verjetnost napake I. vrste večja od željene. Pri željeni 0.01 stopnji značilnosti bo test H_0 zavrnil v približno 0.03 primerih. Pri željeni 0.05 stopnji značilnosti bo test H_0 zavrnil v približno 0.12 primerih. Pri željeni 0.1 stopnji značilnosti bo test H_0 zavrnil v približno 0.2 primerih. To pomeni, da je v vseh primerih test liberalen in posledično nesprejemljiv. Če primerjamo različne vrednosti parametra α opazimo, da se delež zavrnitve testa ne spreminja. Če primerjamo različne velikosti vzorcev opazimo, da se delež zavrnitve testa pravtako ne spreminja. Enako velja tudi za različne vrednosti parametra x_m. Končna ugotovitev je, da velikost testa za dve populaciji porazdeljeni Pareto, kjer smo za izračun p-vrednosti uporabili dve stopinje prostosti nikoli ne zavrne ničelne domneve pri željeni meri, ne glede na velikost vzorca, vrednosti parametra α, vrednosti parametra x_m in stopnjo značilnosti. Ta rezultat je precej presenetljiv, saj smo v teoretičnem delu povedali, da moramo pri χ^2 porazdelitvi upoštevati 2 stopinji prostosti.

Slika 1: Velikost testa - 2 stopinje prostosti.



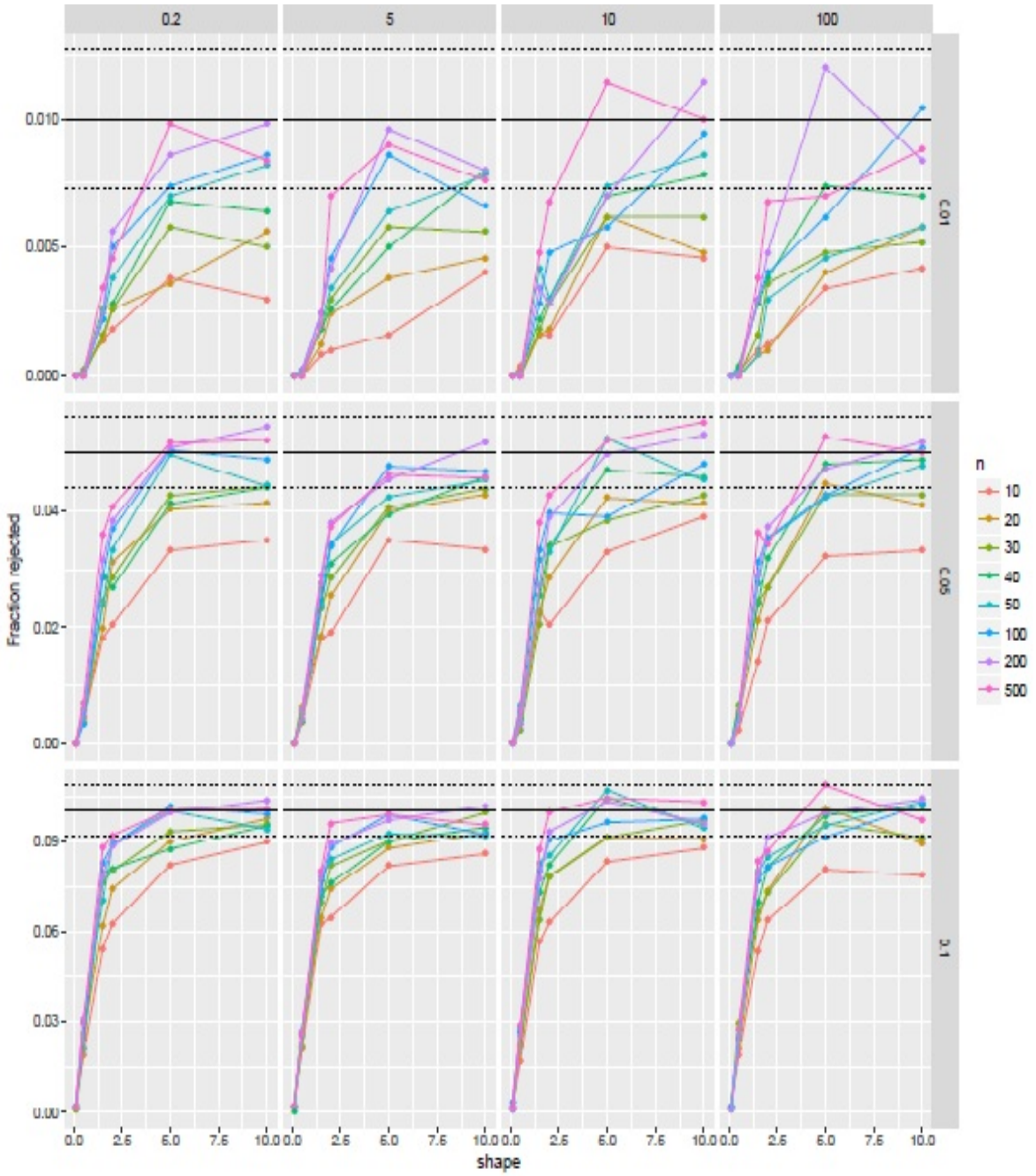
Slika 2: Velikost testa - 3 stopinje prostosti.

Slika 2 prikazuje velikost testa za dve populaciji porazdeljeni Pareto, kjer smo za izračun p-vrednosti uporabili tri stopinje prostosti. Na x-osi so različne vrednosti parametra α(shape). Na y-osi je delež zavrnitve testa. Različne barve krivulj prikazujejo različne velikosti vzorca. Stolpci grafov se razlikujejo po parametru x_m, vrstice grafov pa po željeni stopnji značilnosti. Zaradi vzorčne napake toleriramo manjša odstopanja od željene stopnje značilnosti, zato sprejemo vse točke grafov, ki so znotraj črtkanih vodoravnih črt. Iz slike 2 razberemo, da pri majhnih vzorcih se točke nahajajo izven črtkanih vodoravnih črt, kar pomeni, da je pri majhnih vzorcih dejanska verjetnost napake I. vrste večja od željene. To pomeni, da je v primerih, ko imamo majhen vzorec test liberalen in posledično nesprejemljiv. V primerih, ko imamo velik vzorec, pa se točke večinoma nahajajo znotrah črtkanih vodoravnih črt, kar pomeni, da je test konservativen in posledično sprejemljiv. Če primerjamo različne vrednosti parametra α opazimo, da se delež zavrnitve testa ne spreminja. Enako velja tudi za različne vrednosti parametra x_m. Končna ugotovitev je, da velikost testa za dve populaciji porazdeljeni Pareto, kjer smo za izračun p-vrednosti uporabili tri stopinje prostosti zavrne H_0 pri željeni meri le, če imamo dovolj velik vzorec. Vrednosti parametra α, vrednost parametra x_m in željena stopnja značilnosti na rezultat ne vplivata. Opazimo tudi, da smo s tremi stopinjami prostosti dobili sprejemljiv test, kar smo pričakovali pri testu z dvema stopinjama prostosti.



Slika 3: Velikost testa - permutacijski test.

Slika 3 prikazuje velikost testa za dve populaciji porazdeljeni Pareto z uporabo permutacijskega testa. Na x-osi so različne vrednosti parametra α(shape). Na y-osi je delež zavrnitve testa. Različne barve krivulj prikazujejo različne velikosti vzorca. Stolpci grafov se razlikujejo po parametru x_m, vrstice grafov pa po željeni stopnji značilnosti. Zaradi vzorčne napake toleriramo manjša odstopanja od željene stopnje značilnosti, zato sprejmemo vse točke grafov, ki so znotraj črtkanih vodoravnih črt. Iz slike 3 razberemo, da se večina točk nahaja znotraj črtkanih vodoravnih črt, kar pomeni, da za večino točk velja, da je dejanska verjetnost napake I. vrste v željeni okolici vseh stopenj značilnosti. To pomeni, da je v večini primerih test sprejemljiv. Če primerjamo različne vrednosti parametra α opazimo, da se delež zavrnitve testa ne spreminja. Če primerjamo različne velikosti vzorcev opazimo, da se delež zavrnitve testa pravtako ne spreminja. Enako velja tudi za različne vrednosti parametra x_m. Končna ugotovitev je, da velikost testa za dve populaciji porazdeljeni Pareto z uporabo permutacijskega testa zavrne ničelno domnevo pri željeni meri ne glede na velikost vzorca, velikost parametra α, velikost parametra x_m in stopnjo značilnosti. To dokazuje uporabnost in učinkovitost permutacijskega testa.



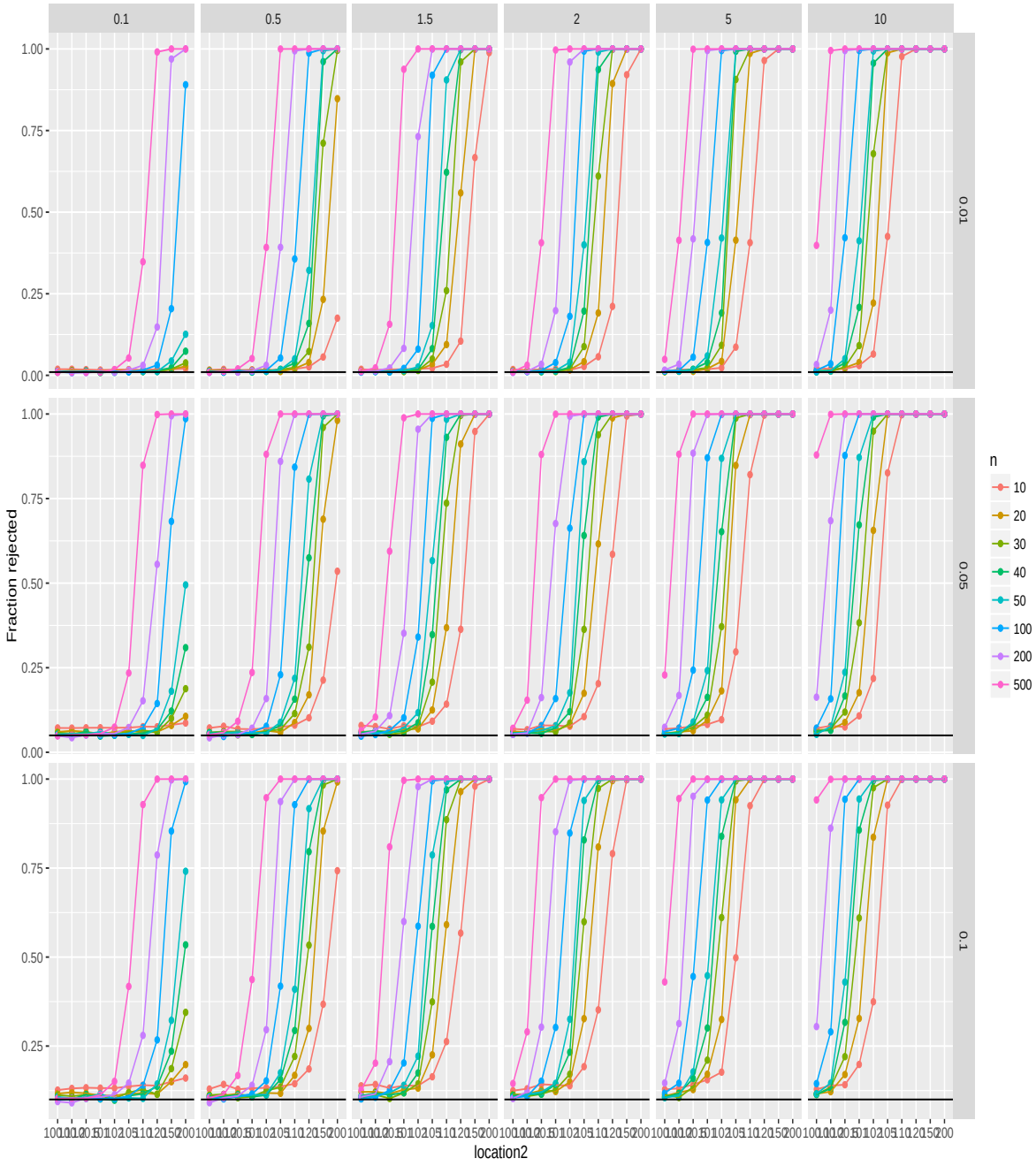
Slika 4: Velikost testa - t-test.

Slika 4 prikazuje velikost testa za dve populaciji porazdeljeni Pareto z uporabo t testa. Na x-osi so različne vrednosti parametra α(shape). Na y-osi je delež zavrnitve testa. Različne barve krivulj prikazujejo različne velikosti vzorca. Stolpci grafov se razlikujejo po parametru x_m, vrstice grafov pa po željeni stopnji značilnosti. Zaradi vzorčne napake toleriramo manjša odstopanja od željene stopnje značilnosti, zato sprejmemo vse točke grafov, ki so znotraj črtkanih vodoravnih črt. Iz slike 4 razberemo, da so nekatere točke znotraj obsega črtkanih vodoravnih črt, nekatere pa pod obsegom ali celo na sami ničli. To pomeni, da v nekaterih primerih je test preveč konservativen. Tudi takšni testi za nas niso sprejemljivi. Opazimo, da ti primeri nastopijo, ko je vrednost parametra α enaka 2.5 ali manj in ko je velikost vzorca manjša od 30. Če primerjamo različne vrednosti parametra x_m opazimo, da se delež zavrnitve testa ne spreminja. Končna ugotovitev je, da velikost testa za dve populaciji porazdeljeni Pareto z uporabo t-testa zavrne ničelno domnevo pri željeni meri, če imamo dovolj velik vzorec in dovolj velik parameter α. Opazimo tudi, da pri večji željeni stopnji značilnosti dobimo bolj zadovoljive teste. Testi pri manjši željeni stopnji značilnosti prevečkrat zavračajo H_0. Manjšo stopnjo značilnosti kot si zadamo, večji vzorec in večji parameter α potrebujemo, da dosežemo željeni test. Velikost parametra x_m na rezultat ne vpliva.

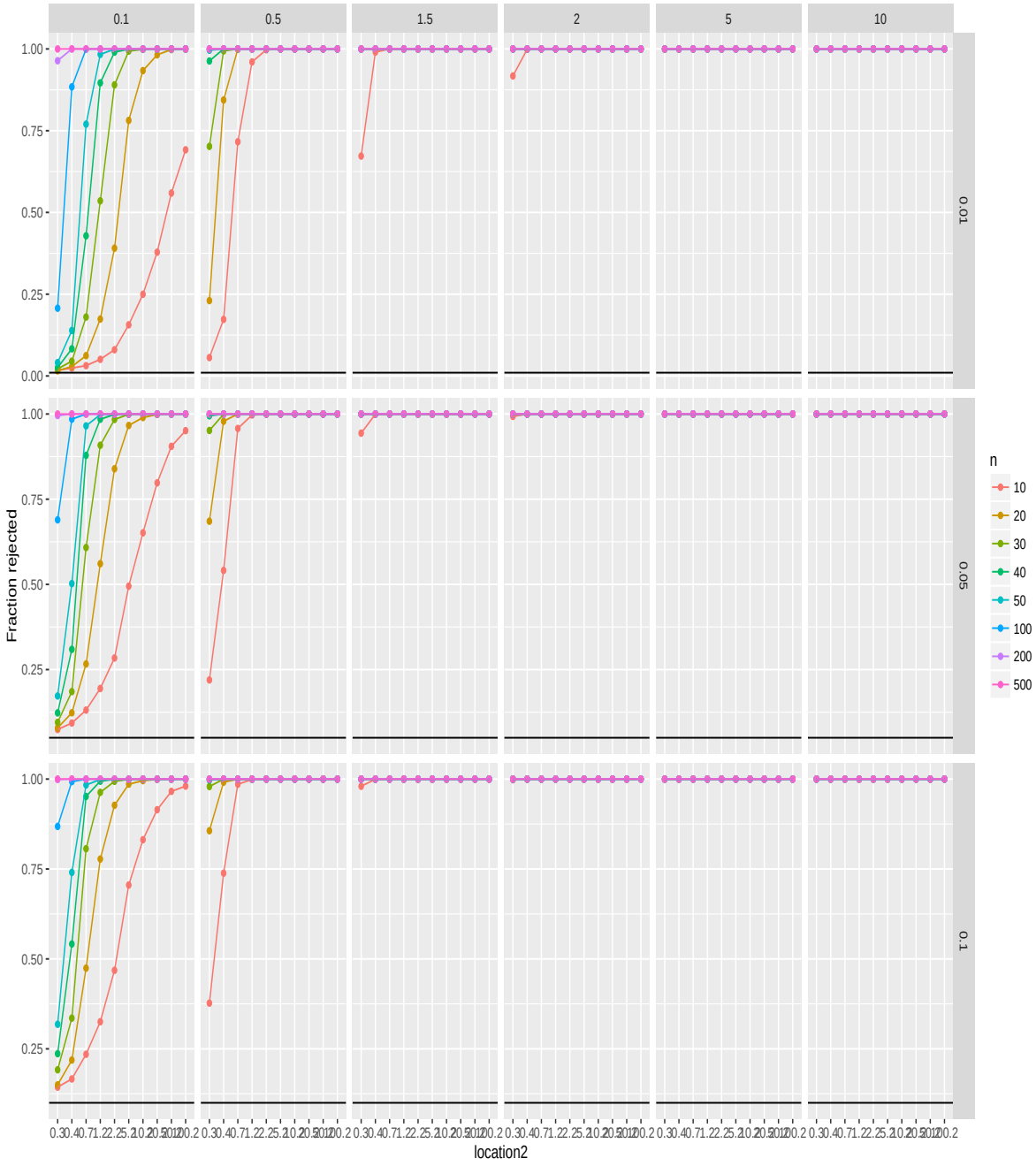
6.2 Moč testov

Slika 5 prikazuje vpliv parametra y_m na moč testa za dve populaciji porazdeljeni Pareto, kjer smo za izračun p-vrednosti uporabili tri stopinje prostosti. Na x-osi so različne vrednosti parametra y_m(location2). Na y-osi je delež zavrnitve testa. Različne barve krivulj prikazujejo različne velikosti vzorca. Stolpci grafov se razlikujejo po parametru α, vrstice grafov pa po željeni stopnji značilnosti. Tu predpostavljamo, da sta parametera α_X in α_Y enaka, vrednost parametra x_m pa je fiksirana pri 100. Slika 6 prikazuje vpliv parametra y_m na moč testa za dve populaciji porazdeljeni Pareto, kjer smo za izračun p-vrednosti uporabili tri stopinje prostosti. Na x-osi so različne vrednosti parametra y_m(location2). Na y-osi je delež zavrnitve testa. Različne barve krivulj prikazujejo različne velikosti vzorca. Stolpci grafov se razlikujejo po parametru α, vrstice grafov pa po željeni stopnji značilnosti. Tu predpostavljamo, da sta parametera α_X in α_Y enaka, vrednost parametra x_m pa je fiksirana pri 0.2. Iz slike 5 razberemo, da so točke na levi strani grafa precej nizko, točke na desni strani grafa pa precej visoko. To nam pove, da bolj kot sta si parametra x_m in y_m različna, večja je moč testa. Če primerjamo grafe po stolpcih opazimo, da je pri desnih stolpcih precej več točk na vrhu grafa, kar nakazuje, da večji kot je parameter α večja je moč testa. Če primerjamo barve posameznih krivulj opazimo, da večji vzorec kot krivulja prikazuje, več točk ima

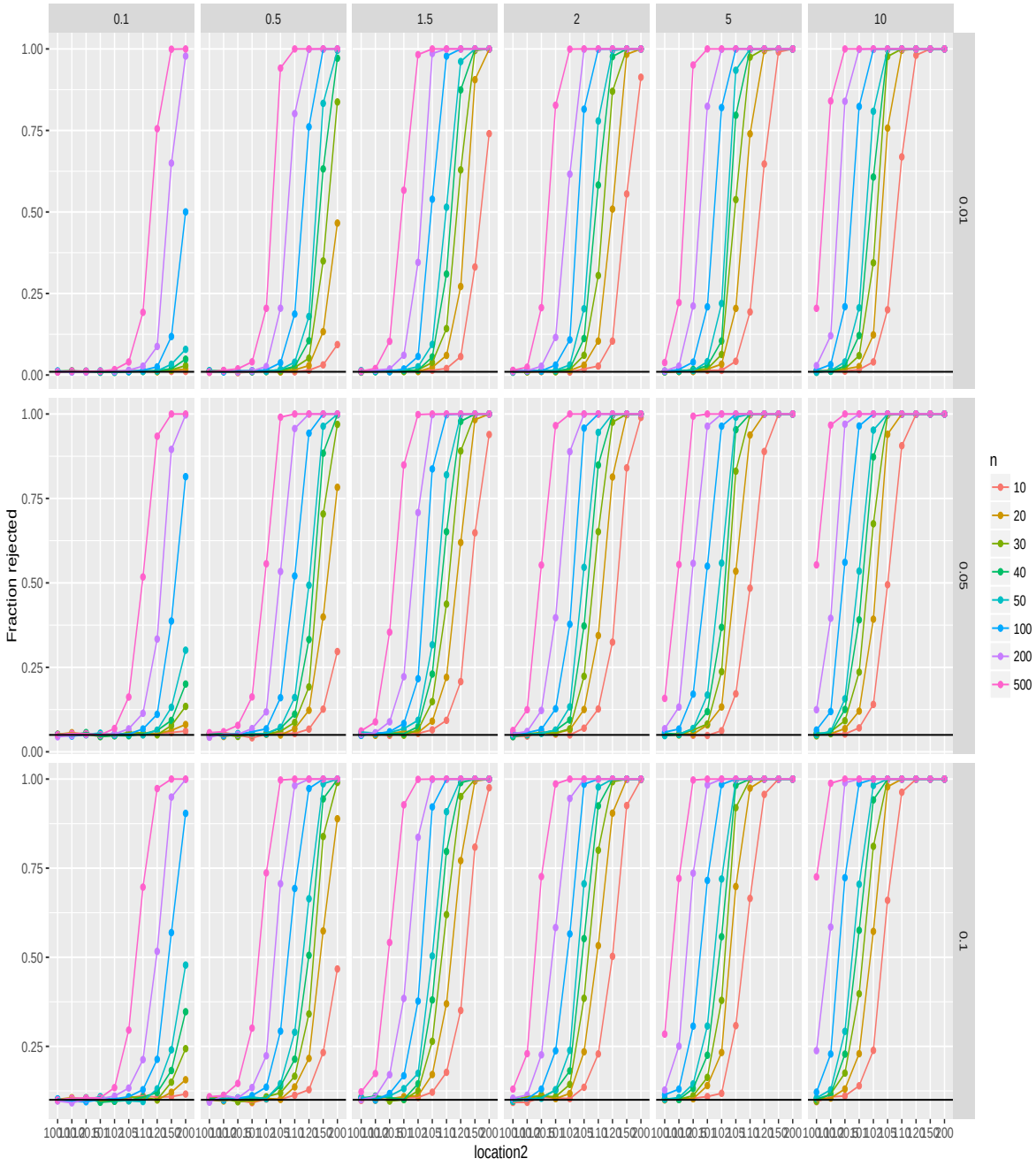
pri vrhu grafa. To nam pove, da večji kot je vzorec, večja je moč testa. Če primerjamo sliki 5 in 6 ugotovimo, da na slednji imamo precej več točk na vrhu grafa, kar pomeni, da ima precejšen vpliv tudi odstotkovna razlika med parametroma x_m in y_m. Večja kot je odstotkovna razlika, večja je moč testa. Končna ugotovitev je, da bolj kot sta si x_m in y_m različna, večja je moč testa, kar smo tudi pričakovali. Velik vpliv ima tudi odstotkovna razlika med x_m in y. Opazimo tudi, da večja kot sta vrednost parametera α in velikost vzorca, večja je moč testa, medtem, ko željena stopnja značilnosti na razultat ne vpliva.



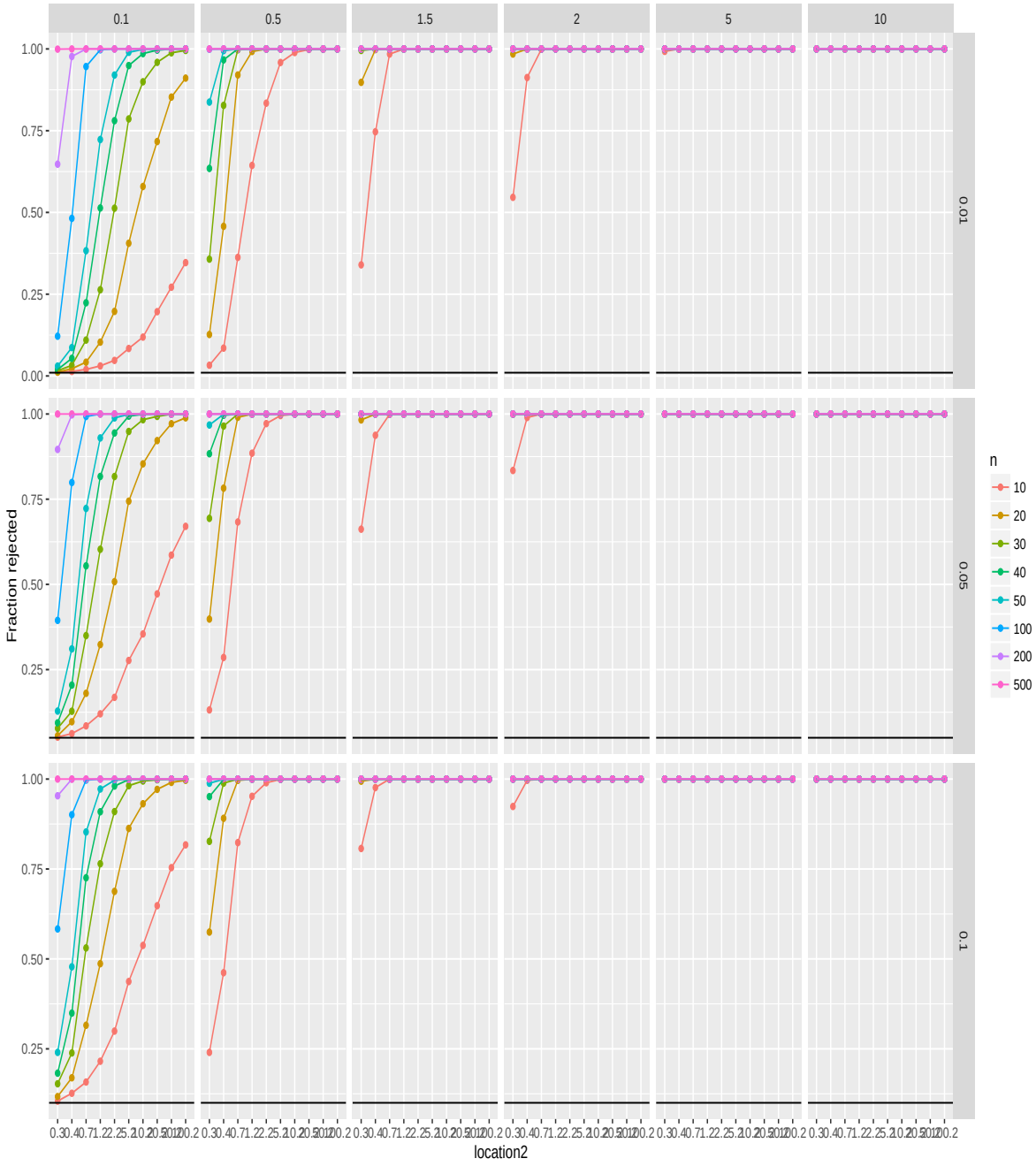
Slika 5: Moč testa za različna x_m, $x_m = 100$ - 3 stopinje prostosti.



Slika 6: Moč testa za različna x_m, $x_m = 0.2$ - 3 stopinje prostosti.

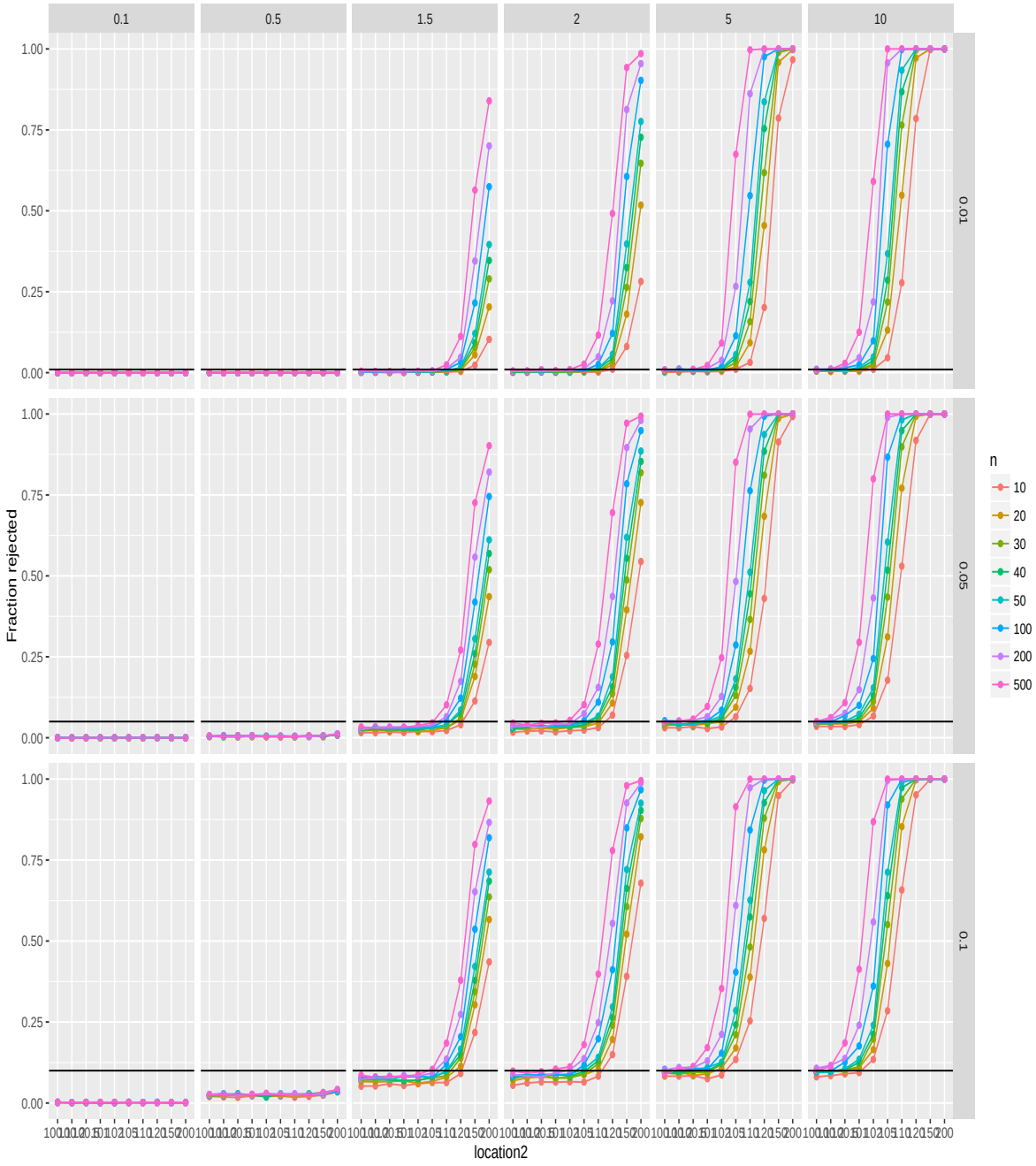


Slika 7: Moč testa za različna x_m, $x_m = 100$ - permutacijski test.

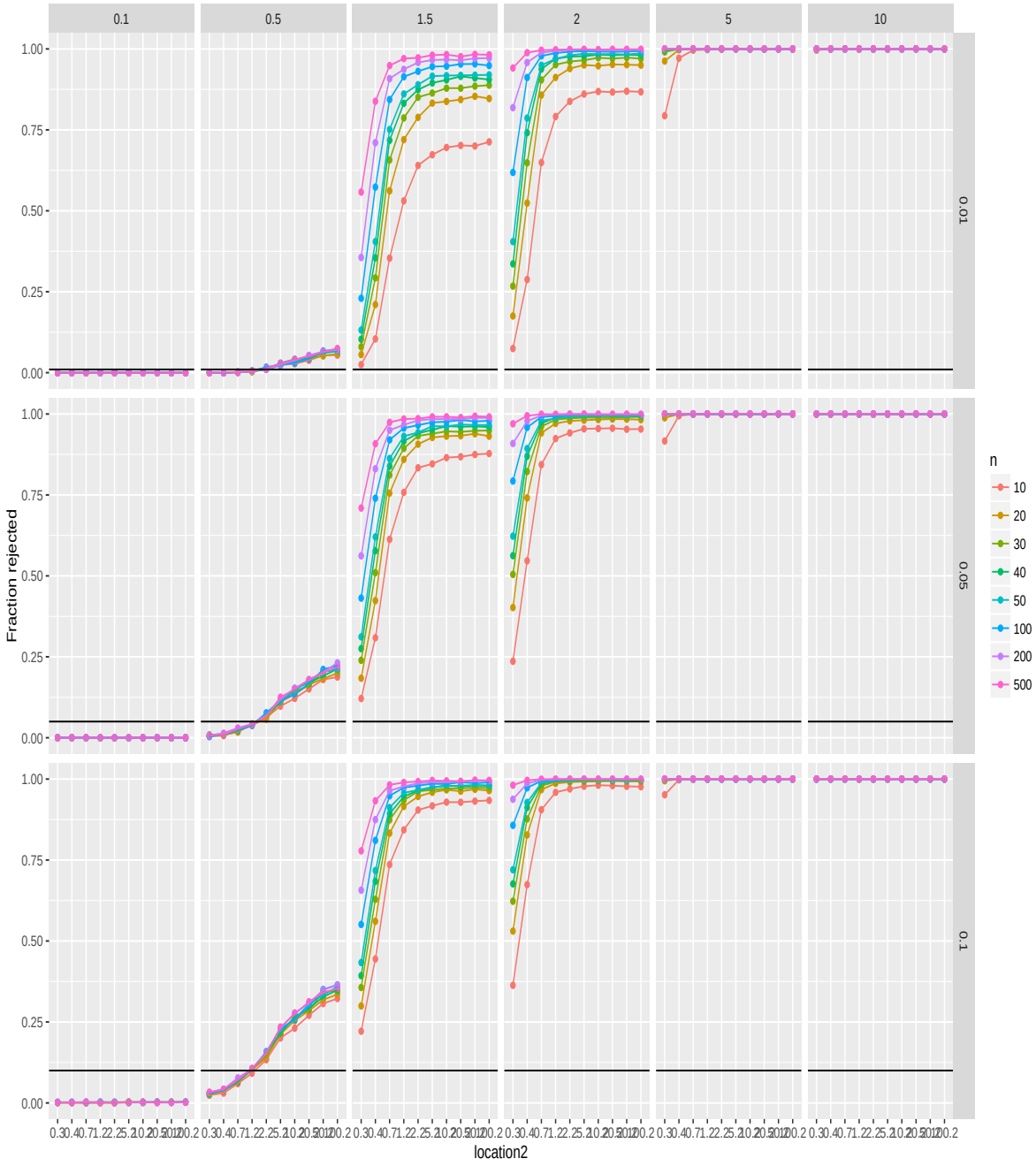


Slika 8: Moč testa za različna x_m, $x_m = 0.2$ - permutacijski test.

Slika 7 prikazuje vpliv parametra y_m na moč testa za dve populaciji porazdeljeni Pareto s permutacijskim testom. Na x-osi so različne vrednosti parametra y_m(location2). Na y-osi je delež zavrnitve testa. Različne barve krivulj prikazujejo različne velikosti vzorca. Stolpci grafov se razlikujejo po parametru α, vrstice grafov pa po željeni stopnji značilnosti. Tu predpostavljamo, da sta parametera α_X in α_Y enaka, vrednost parametra x_m pa je fiksirana pri 100. Slika 8 prikazuje vpliv parametra y_m na moč testa za dve populaciji porazdeljeni Pareto s permutacijskim testom. Na x-osi so različne vrednosti parametra y_m(location2). Na y-osi je delež zavrnitve testa. Različne barve krivulj prikazujejo različne velikosti vzorca. Stolpci grafov se razlikujejo po parametru α, vrstice grafov pa po željeni stopnji značilnosti. Tu predpostavljamo, da sta parametera α_X in α_Y enaka, vrednost parametra x_m pa je fiksirana pri 0.2. Iz slike 7 razberemo, da so točke na levi strani grafa precej nizko, točke na desni strani grafa pa precej visoko. To nam pove, da bolj kot sta si parametra x_m in y_m različna, večja je moč testa. Če primerjamo grafe po stolpcih opazimo, da je pri desnih stolpcih precej več točk na vrhu grafa, kar nakazuje, da večji kot je parameter α večja je moč testa. Če primerjamo barve posameznih krivulj opazimo, da večji vzorec kot krivulja prikazuje, več točk ima pri vrhu grafa. To nam pove, da večji kot je vzorec, večja je moč testa. Manjši vpliv ima tudi željena stopnja značilnosti. Pri večji stopnji značilnosti krivulje hitreje dosežejo vrh grafa, kar pomeni, da večja kot je željena stopnja značilnosti večja je moč testa. Če primerjamo sliki 7 in 8 ugotovimo, da na slednji imamo precej več točk na vrhu grafa, kar pomeni, da ima precejšen vpliv tudi odstotkovna razlika med parametroma x_m in y_m. Večja kot je odstotkovna razlika, večja je moč testa. Končna ugotovitev je, da bolj kot sta si x_m in y_m različna, večja je moč testa, kar smo tudi pričakovali. Velik vpliv ima tudi odstotkovna razlika med x_m in y. Opazimo tudi, da večja kot sta vrednost parametera α, velikost vzorca in željena stopnja značilnosti, večja je moč testa.

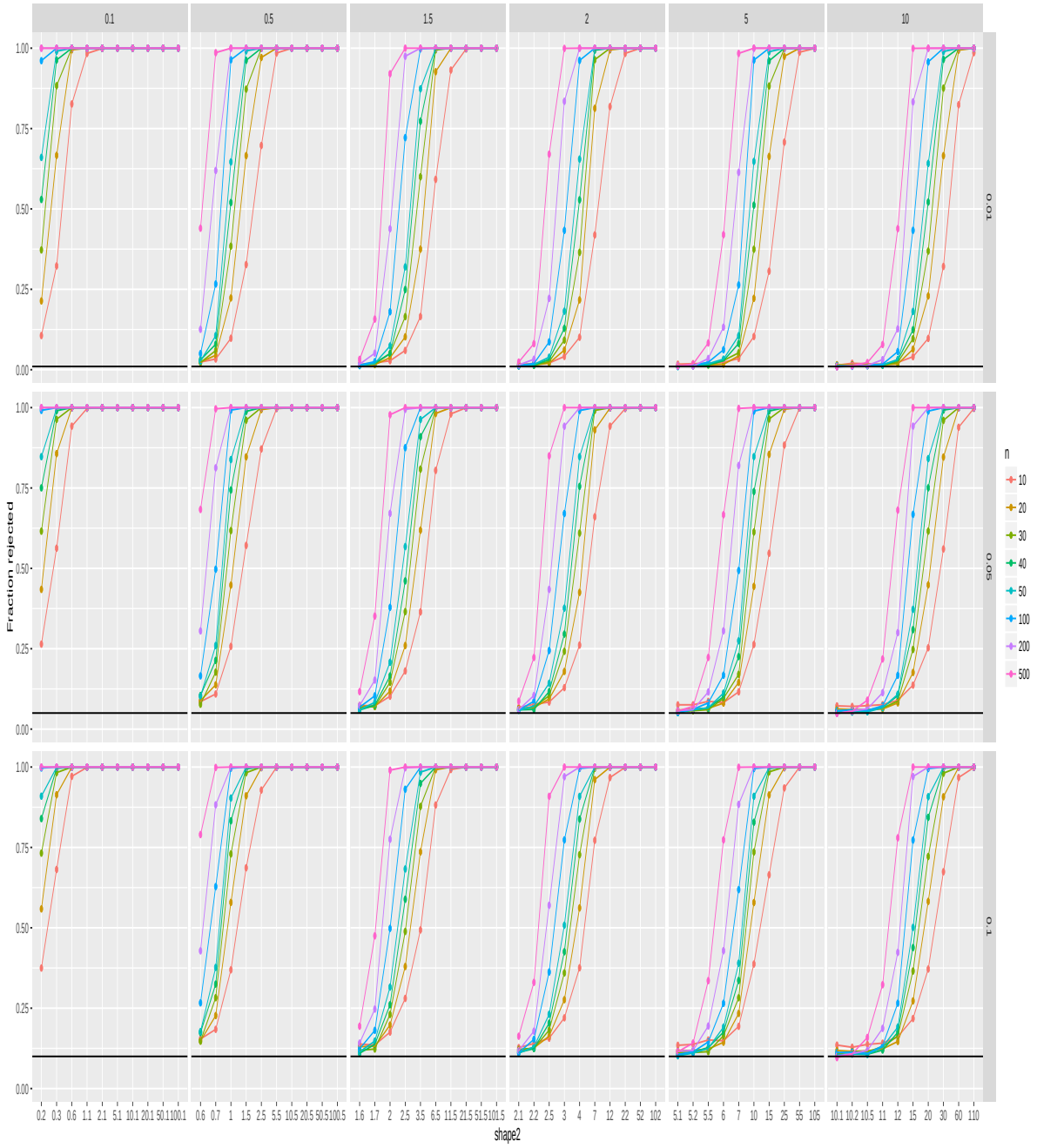


Slika 9: Moč testa za različna x_m, $x_m = 100$ - t-test.

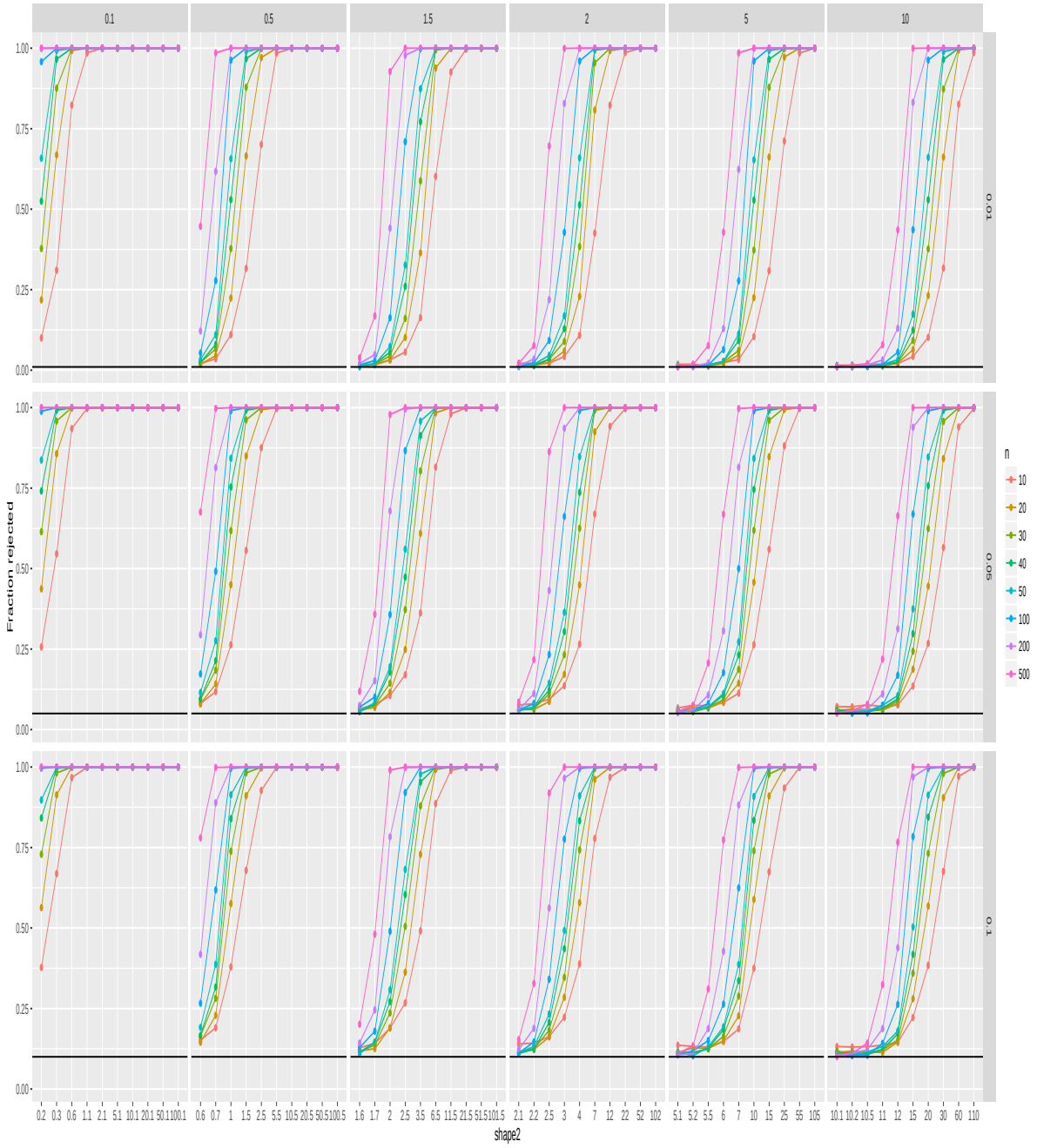


Slika 10: Moč testa za različna x_m, $x_m = 0.2$ - t-test.

Slika 9 prikazuje vpliv parametra y_m na moč testa za dve populaciji porazdeljeni Pareto s t-testom. Na x-osi so različne vrednosti parametra y_m(location2). Na y-osi je delež zavrnitve testa. Različne barve krivulj prikazujejo različne velikosti vzorca. Stolpci grafov se razlikujejo po parametru α, vrstice grafov pa po željeni stopnji značilnosti. Tu predpostavljamo, da sta parametera α_X in α_Y enaka, vrednost parametra x_m pa je fiksirana pri 100. Slika 10 prikazuje vpliv parametra y_m na moč testa za dve populaciji porazdeljeni Pareto s t-testom. Na x-osi so različne vrednosti parametra y_m(location2). Na y-osi je delež zavrnitve testa. Različne barve krivulj prikazujejo različne velikosti vzorca. Stolpci grafov se razlikujejo po parametru α, vrstice grafov pa po željeni stopnji značilnosti. Tu predpostavljamo, da sta parametera α_X in α_Y enaka, vrednost parametra x_m pa je fiksirana pri 0.2. Iz slike 9 razberemo predvsem to, da pri vrednosti parametra α manjši od 1.5 zelo skromno moč testa, ne glede na ostale parametre. Vidimo, da so točke na levi strani grafa precej nizko, točke na desni strani grafa pa precej visoko, kar nam pove, da večja kot je razlika med parametroma x_m in y_m, večja je moč testa. Opazimo tudi, da večji vzorec kot krivulja prikazuje, več točk ima pri vrhu grafa. To nam pove, da večji kot je vzorec, večja je moč testa. Vendar v primerjavi s prejšnjimi testi velikost vzorca nima takšnega vpliva, saj so krivulje bolj skupaj kot pri testih za različna x_m. Če primerjamo različne vrednosti željenih stopenj značilnosti vidimo, da so krivulje skoraj identične. To nakazuje, da željena stopnja značilnosti nima vpliva na rezultat. Če primerjamo sliki 9 in 10 opazimo, da slednja ima precej več točk na vrhu grafa, kar pomeni, da ima precejšen vpliv tudi odstotkovna razlika med parametroma x_m in y_m. Večja kot je odstotkovna razlika med njima, večja je moč testa. Končna ugotovitev je, da bolj kot sta si x_m in y_m različna, večja je moč testa. Velik vpliv ima tudi odstotkovna razlika med x_m in y. Opazimo tudi, da večja kot sta vrednost parametera α in velikost vzorca, večja je moč testa, medtem, ko željena stopnja značilnosti na razultat ne vpliva. Za majhne vrednosti parametra α t-test ne zagotavlja željene moči, ne glede na ostale dejavnike.

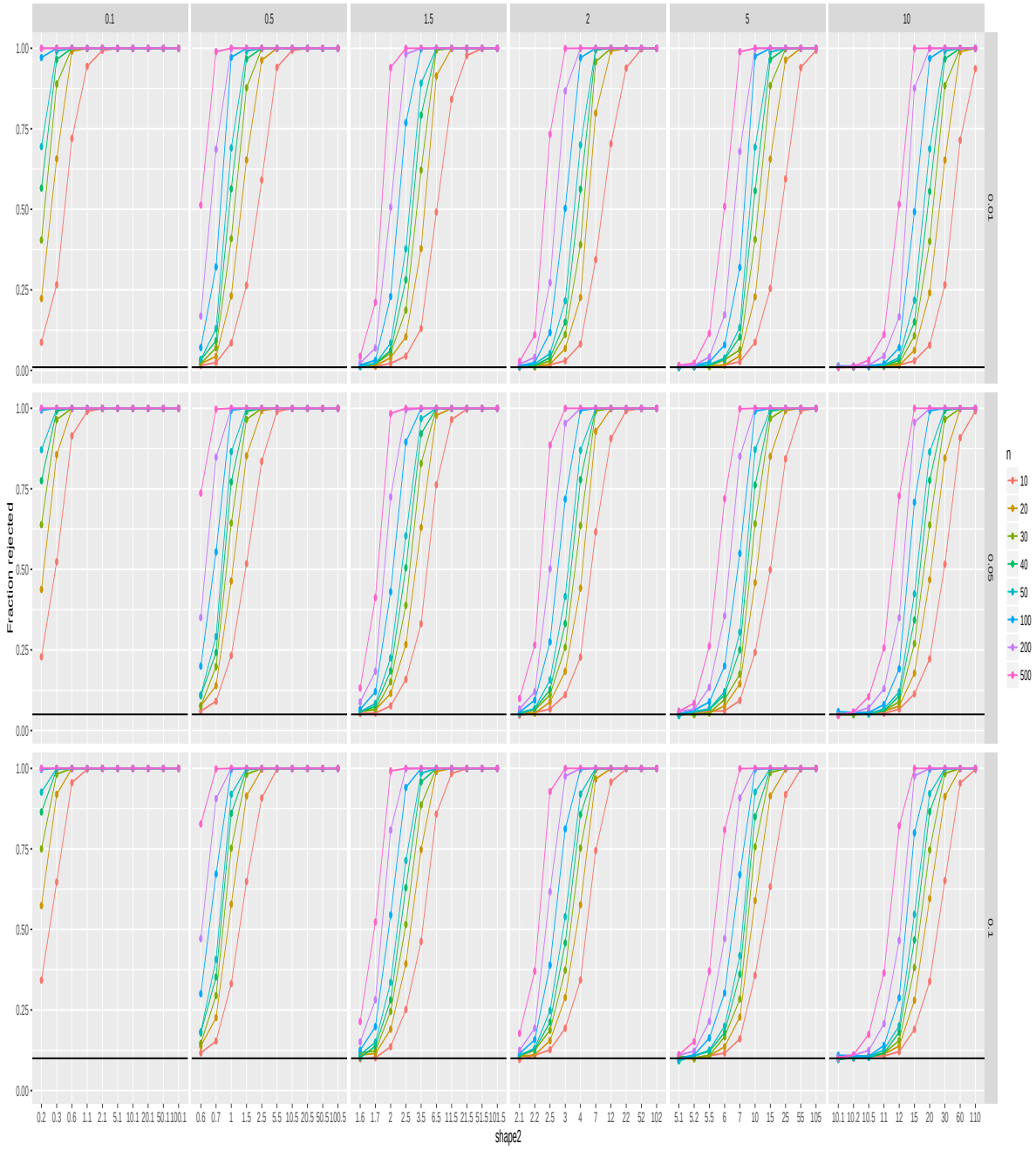


Slika 11: Moč testa za različna α, $x_m = 100$ - 3 stopinje prostosti.

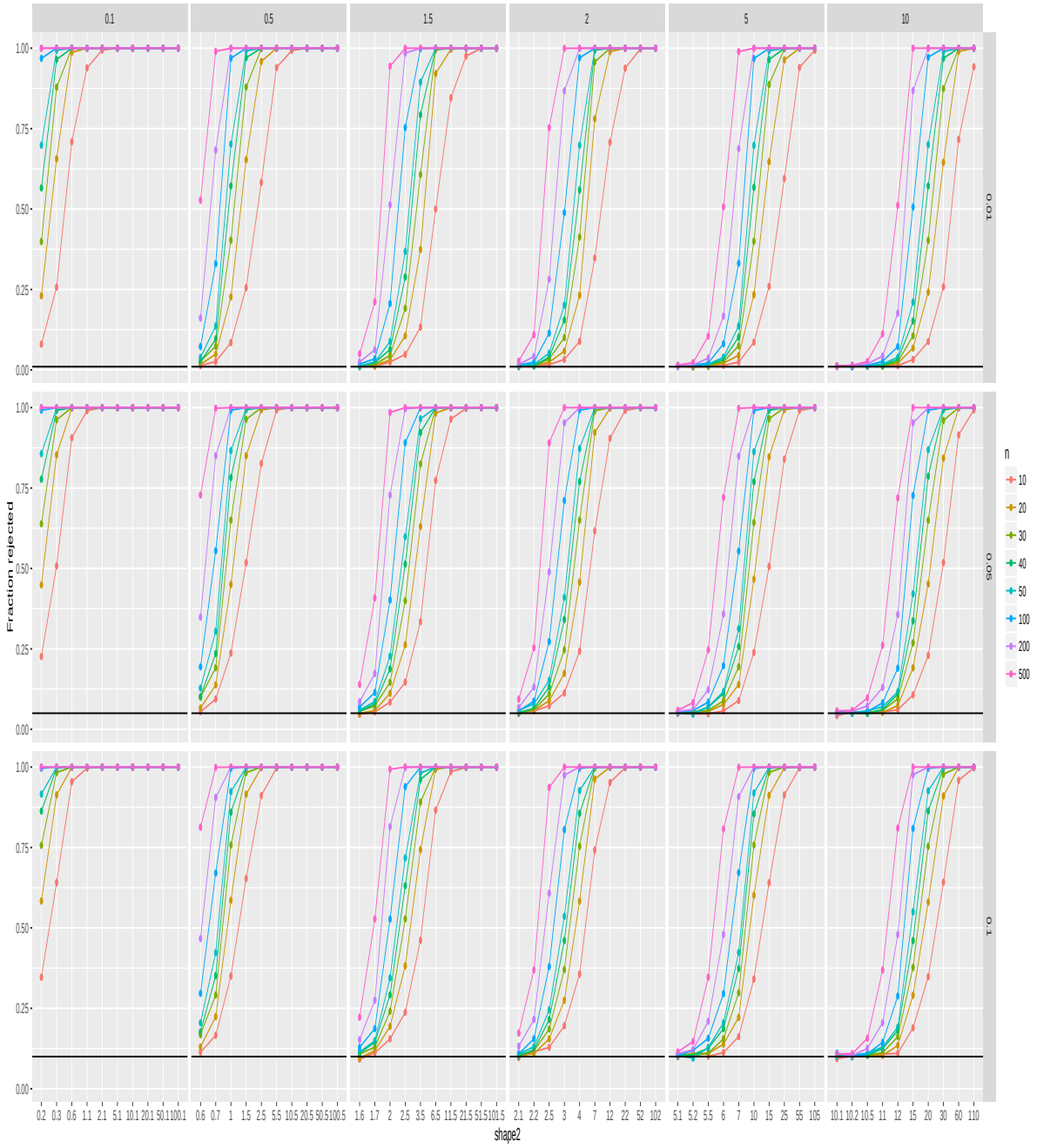


Slika 12: Moč testa za različna α, $x_m = 0.2$ - 3 stopinje prostosti.

Slika 11 prikazuje vpliv parametra α na moč testa za dve populaciji porazdeljeni Pareto, kjer smo za izračun p-vrednosti uporabili tri stopinje prostosti. Na x-osi so različne vrednosti parametra α_Y(shape2). Na y-osi je delež zavrnitve testa. Različne barve krivulj prikazujejo različne velikosti vzorca. Stolpci grafov se razlikujejo po parametru α_X, vrstice grafov pa po željeni stopnji značilnosti. Tu predpostavljamo, da sta parametera x_m in y_m enaka in fiksirana pri 100. Slika 12 prikazuje vpliv parametra α na moč testa za dve populaciji porazdeljeni Pareto, kjer smo za izračun p-vrednosti uporabili tri stopinje prostosti. Na x-osi so različne vrednosti parametra α_Y(shape2). Na y-osi je delež zavrnitve testa. Različne barve krivulj prikazujejo različne velikosti vzorca. Stolpci grafov se razlikujejo po parametru α_X, vrstice grafov pa po željeni stopnji značilnosti. Tu predpostavljamo, da sta parametera x_m in y_m enaka in fiksirana pri 0.2. Iz slike 11 razberemo, da so točke na levi strani grafa precej nizko, točke na desni strani grafa pa precej visoko. To nam pove, da bolj kot sta si parametra α_X in α_Y različna, večja je moč testa. Če primerjamo grafe po stolpcih opazimo, da je pri levih stolpcih precej več točk na vrhu grafa, kar nakazuje, da večja kot je odstotkovna razlika med parametroma α_X in α_Y, večja je moč testa. Opazimo, da večji vzorec kot krivulja prikazuje, več točk ima pri vrhu grafa. To nam pove, da večji kot je vzorec, večja je moč testa. Tudi tu so krivulje različnih barv precej skupaj, zato lahko trdimo, da velikost vzorca nima takšnega vpliva v primerjavi s testi za različna x_m. Če primerjamo različne vrednosti željenih stopenj značilnosti vidimo, da so krivulje skoraj identične, kar pomeni, da željena stopnja značilnosti nima vpliva na rezultat. Če primerjamo sliki 11 in 12 vidimo, da med grafi ni razlik, kar nam pove, da različne vrednosti parametra x_m na rezultat nimajo vpliva. Končna ugotovitev je, da bolj kot sta si α_X in α_Y različna, večja je moč testa, kar smo v teoretičnem delu tudi trdili. Velik vpliv ima tudi odstotkovna razlika med α_X in α_Y. Večja kot je odstotkovna razlika, večja je moč testa. Manjši vpliv ima tudi velikost vzorca, medtem, ko vrednost parametra x_m in željena stopnja značilnosti na razultat nimata vpliva.

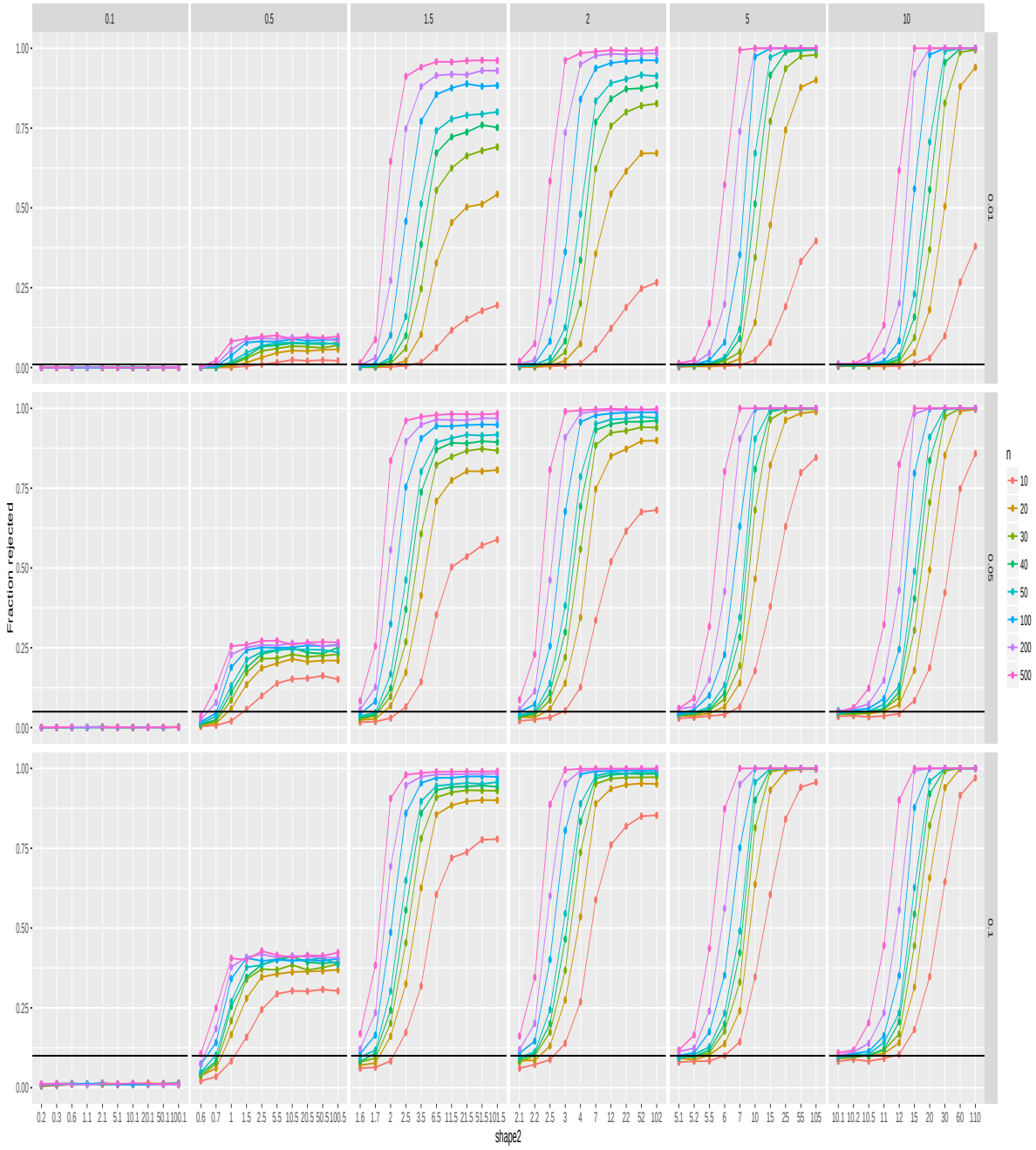


Slika 13: Moč testa za različna α, $x_m = 100$ - permutacijski test.

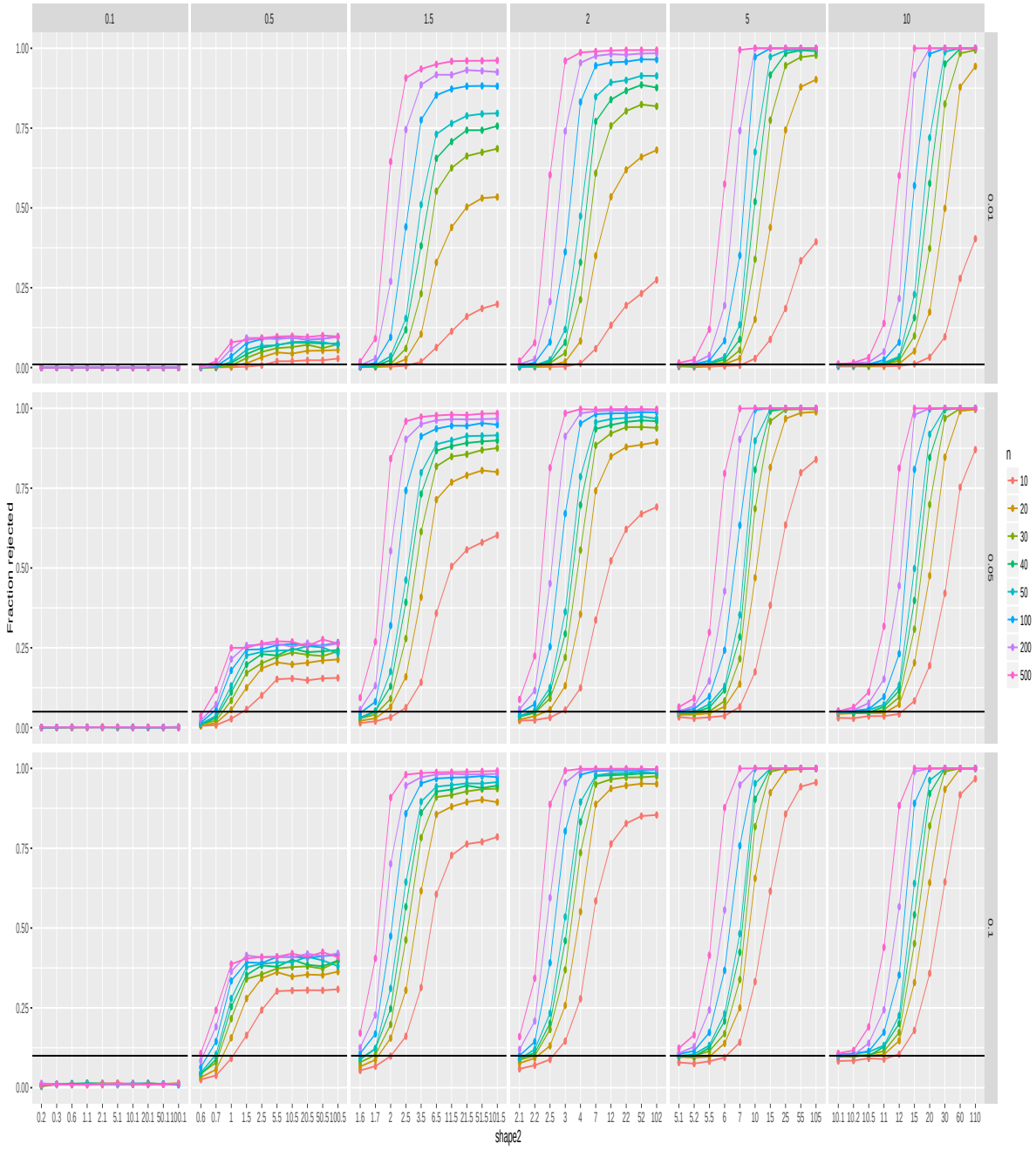


Slika 14: Moč testa za različna α, $x_m = 0.2$ - permutacijski test.

Slika 13 prikazuje vpliv parametra α na moč testa za dve populaciji porazdeljeni Pareto s permutacijskim testom. Na x-osi so različne vrednosti parametra α_Y(shape2). Na y-osi je delež zavrnitve testa. Različne barve krivulj prikazujejo različne velikosti vzorca. Stolpci grafov se razlikujejo po parametru α_X, vrstice grafov pa po željeni stopnji značilnosti. Tu predpostavljamo, da sta vrednosti parametera x_m in y_m enaki in fiksirani pri 100. Slika 14 prikazuje vpliv parametra α na moč testa za dve populaciji porazdeljeni Pareto s permutacijskim testom. Na x-osi so različne vrednosti parametra α_Y(shape2). Na y-osi je delež zavrnitve testa. Različne barve krivulj prikazujejo različne velikosti vzorca. Stolpci grafov se razlikujejo po parametru α_X, vrstice grafov pa po željeni stopnji značilnosti. Tu predpostavljamo, da sta vrednosti parametera x_m in y_m enaki in fiksirani pri 0.2. Iz slike 13 razberemo, da so točke na levi strani grafa precej nizko, točke na desni strani grafa pa precej visoko, kar nam nakazuje, da bolj kot sta si parametra α_X in α_Y različna, večja je moč testa. Če primerjamo grafe po stolpcih opazimo, da je pri levih stolpcih precej več točk na vrhu grafa, kar pomeni, da večja kot je odstotkovna razlika med parametroma α_X in α_Y, večja je moč testa. Vidimo, da večji vzorec kot krivulja prikazuje, več točk ima pri vrhu grafa. To nakazuje, da večji kot je vzorec, večja je moč testa. Tudi tu so krivulje različnih barv precej skupaj, zato lahko trdimo, da velikost vzorca nima takšnega vpliva v primerjavi s testi za različna x_m. Če primerjamo različne vrednosti željenih stopenj značilnosti vidimo, da so krivulje skoraj identične, kar pomeni, da željena stopnja značilnosti nima vpliva na rezultat. Če primerjamo sliki 13 in 14 opazimo, da med grafi ni razlik, kar nam pove, da različne vrednosti parametra x_m na rezultat nimajo vpliva. Končna ugotovitev je, da bolj kot sta si α_X in α_Y različna, večja je moč testa, kar smo v teoretičnem delu tudi trdili. Velik vpliv ima tudi odstotkovna razlika med α_X in α_Y. Večja kot je odstotkovna razlika, večja je moč testa. Manjši vpliv ima tudi velikost vzorca, medtem, ko vrednost parametra x_m in željena stopnja značilnosti na razultat nimata vpliva.



Slika 15: Moč testa za različna α, $x_m = 100$ - t-test.



Slika 16: Moč testa za različna α, $x_m = 0.2$ - t-test.

Slika 15 prikazuje vpliv parametra α na moč testa za dve populaciji porazdeljeni Pareto s t-testom. Na x-osi so različne vrednosti parametra α_Y(shape2). Na y-osi je delež zavrnitve testa. Različne barve krivulj prikazujejo različne velikosti vzorca. Stolpci grafov se razlikujejo po parametru α_X, vrstice grafov pa po željeni stopnji značilnosti. Tu predpostavljamo, da sta vrednosti parametera x_m in y_m enaki in fiksirani pri 100. Slika 16 prikazuje vpliv parametra α na moč testa za dve populaciji porazdeljeni Pareto s t-testom. Na x-osi so različne vrednosti parametra α_Y(shape2). Na y-osi je delež zavrnitve testa. Različne barve grafov prikazujejo različne velikosti vzorca. Stolpci grafov se razlikujejo po parametru α_X, vrstice grafov pa po željeni stopnji značilnosti. Tu predpostavljamo, da sta vrednosti parametera x_m in y_m enaki in fiksirani pri 0.2. Iz slike 15 najprej razberemo, da pri vrednosti parametrov α_X in α_Y manjši od 1.5 dobimo precej skromno moč testa, ne glede na ostale parametre. Točke na levi strani grafa so precej nizko, točke na desni strani grafa pa precej visoko, kar nam pove, da bolj kot sta si parametra α_X in α_Y različna, večja je moč testa. Vidimo, da večji vzorec kot krivulja prikazuje, več točk ima pri vrhu grafa. To nakazuje, da večji kot je vzorec, večja je moč testa. Tu so krivulje različnih barv nekoliko narazen, zato lahko trdimo, da velikost vzorca ima večji vpliv v primerjavi s prejšnjimi testi za različna α. Če primerjamo različne vrednosti željenih stopenj značilnosti vidimo, da se krivulje le delno spremenijo, kar pomeni, da željena stopnja značilnosti ima zelo majhen vpliv na rezultat. Če primerjamo sliki 15 in 16 opazimo, da med grafi ni razlik, kar nam pove, da različne vrednosti parametra x_m na rezultat nimajo vpliva. Končna ugotovitev je, da bolj kot sta si α_X in α_Y različna, večja je moč testa, kar smo v teoretičnem delu tudi trdili. Vpliv na rezultat ima tudi velikost vzorca. Odstotkov razlika med parametroma α_X in α_Y, vrednost parametra x_m in željena stopnja značilnosti na razultat nimajo vpliva.

7 Zaključek

V zaključni nalogi smo skušali z uporabo permutacijskega testa preveriti enakost povprečij dveh populacij porazdeljeni Pareto. Ugotovili smo, da je permutacijski test zelo dober način preverjanja statističnih domnev, saj smo z njegovo uporabo dobili zelo dobre rezultate. Ni pa to edini test, s katerim smo prišli do željenih rezultatov. Test razmerja verjetij nam je najboljše rezultate dal pri treh stopinjah prostosti. Pri t-testu je precej vpliva imel parameter α. Za dovolj velik α in pri dovolj velikem vzorcu smo tudi pri t-testu dobili željene rezultate.

8 Literatura

[1] J.L. Devore, N.R. Farnum in Applied statistics for engineers and scientists, *Thomson Brooks/Cole, Second Edition, 2005. (Ni citirano.)*

Devore

[2] P. Legendre in L. Legendre, Statistical testing by permutation. *1998. Numerical ecology, 2nd English edition. Elsevier Science BV, Amsterdam, http://biol09.biol.umontreal.ca/PLcourses/Statistical_tests.pdf*, sprejeto v objavo. *(Citirano na strani 14.)*

[3] J.A. Rice, *Mathematical Statistics and Data Analysis, Third Edition*, University of California, Berkeley, 2006. *(Citirano na straneh 3, 4, 6, 7 in 9.)*

[4] S.S. Wilks, The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *The Annals of Mathematical Statistics, vol. 9, no. 1, 1938, pp. 60–62. JSTOR, www.jstor.org/stable/2957648*, sprejeto v objavo. *(Citirano na strani 4.)*

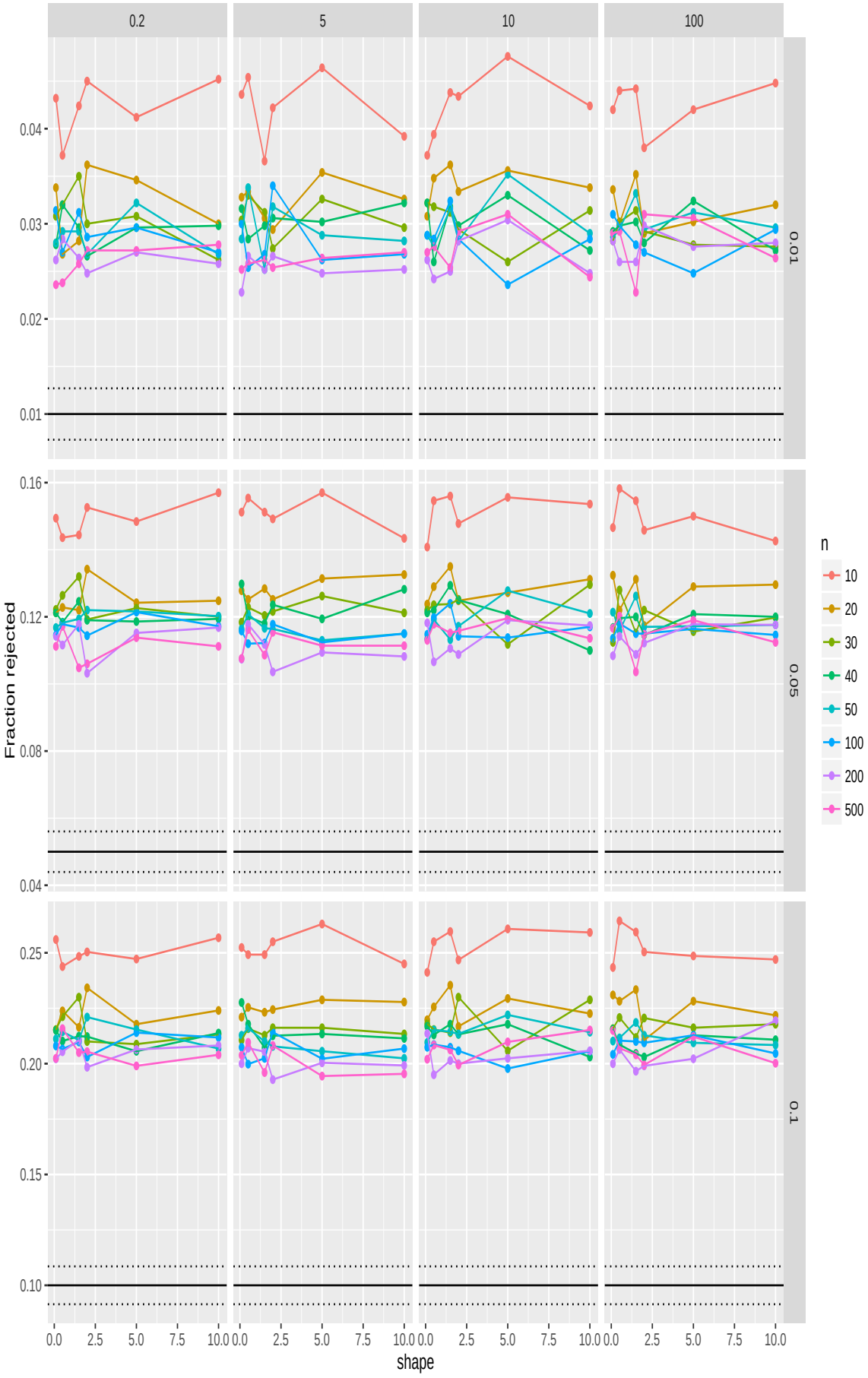
[5] *Normal distribution*,
`https://en.wikipedia.org/wiki/Normal_distribution`. (Datum ogleda: 3. 8. 2017.) *(Citirano na strani 5.)*

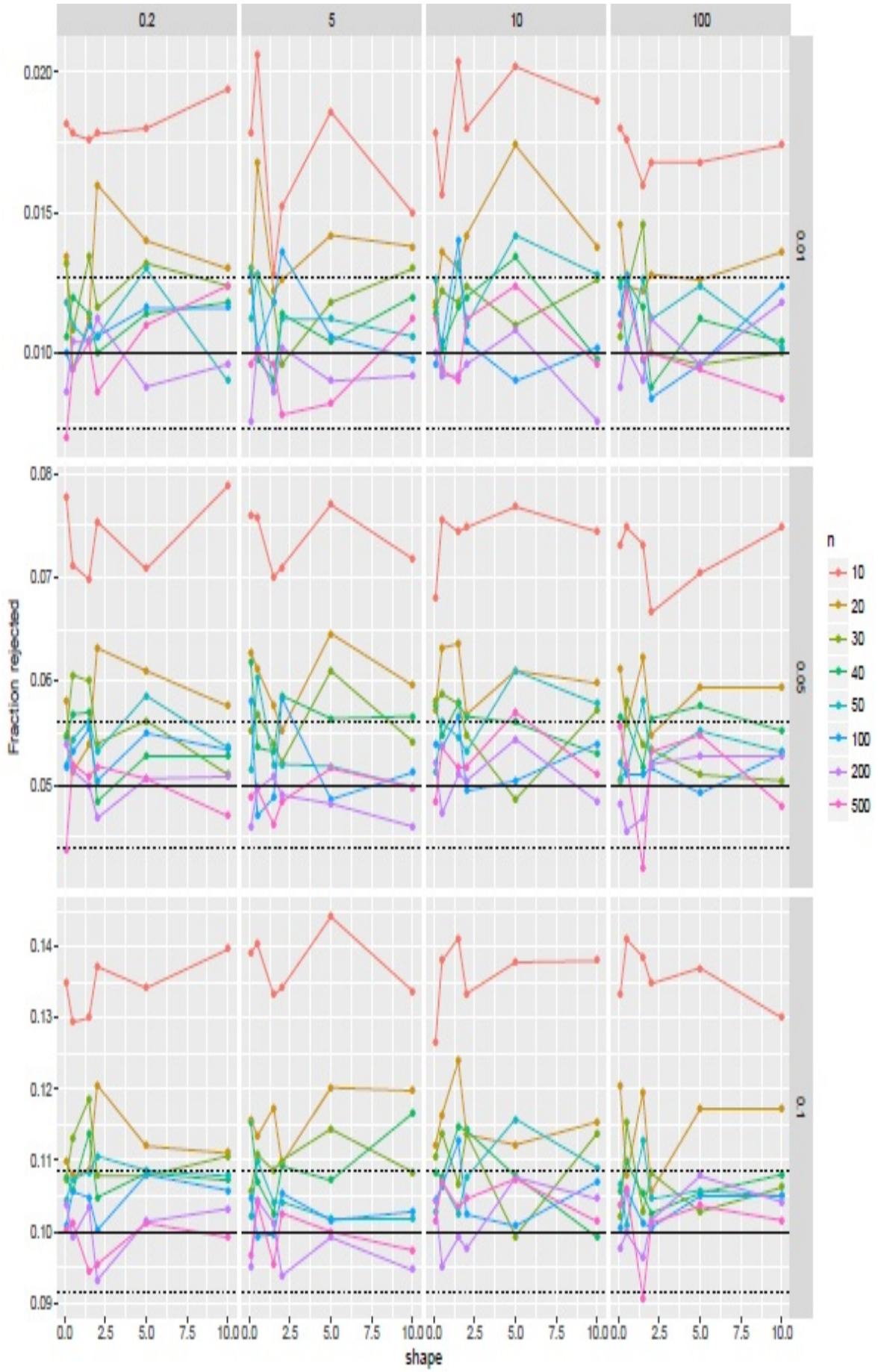
[6] *What is R?*,
`https://www.r-project.org/about.html`. (Datum ogleda: 14. 6. 2017.) *(Citirano na strani 16.)*

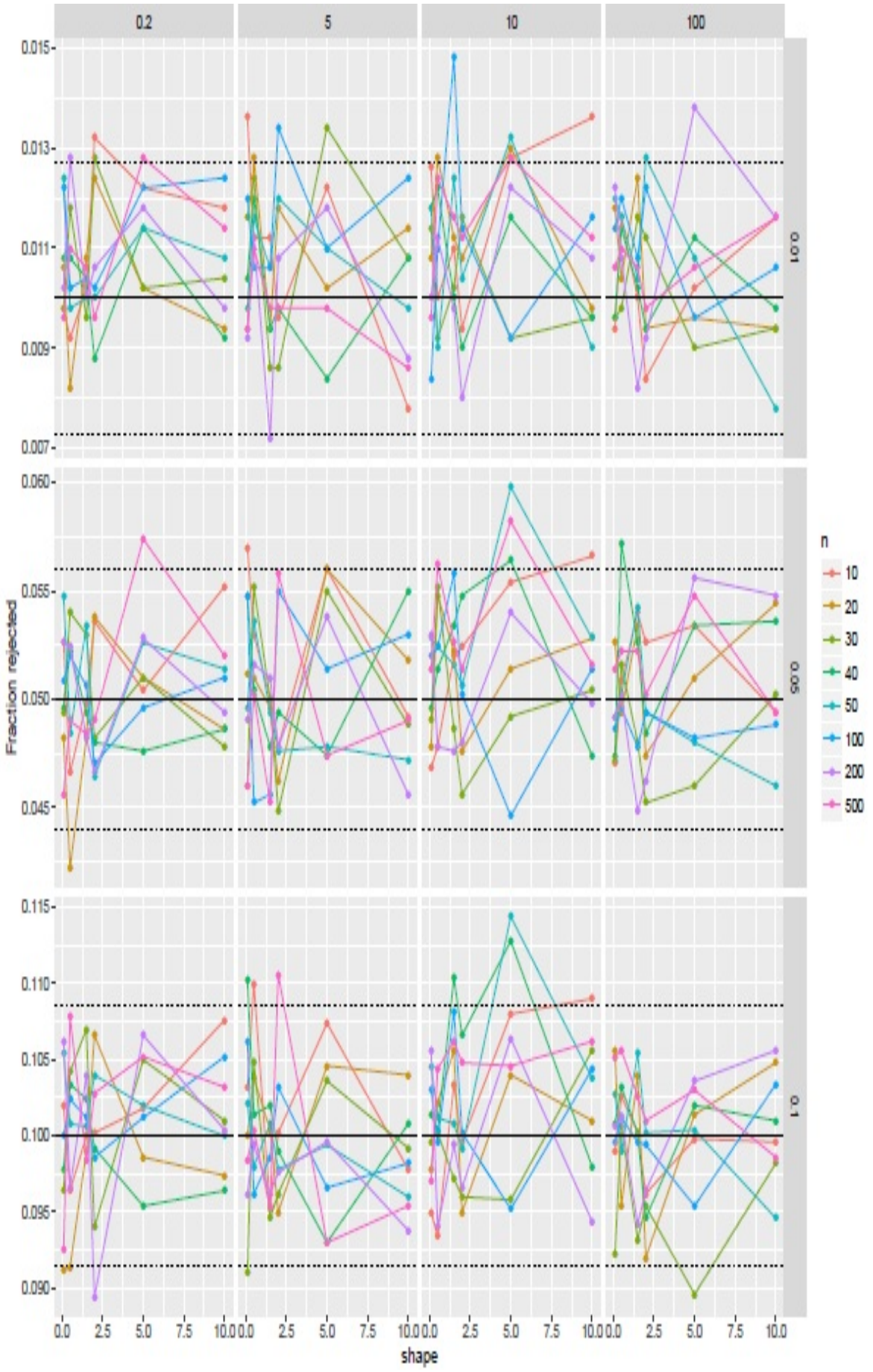
[7] *Pareto distribution*,
`https://en.wikipedia.org/wiki/Pareto_distribution`. (Datum ogleda: 22. 5. 2017.) *(Citirano na straneh 10 in 12.)*

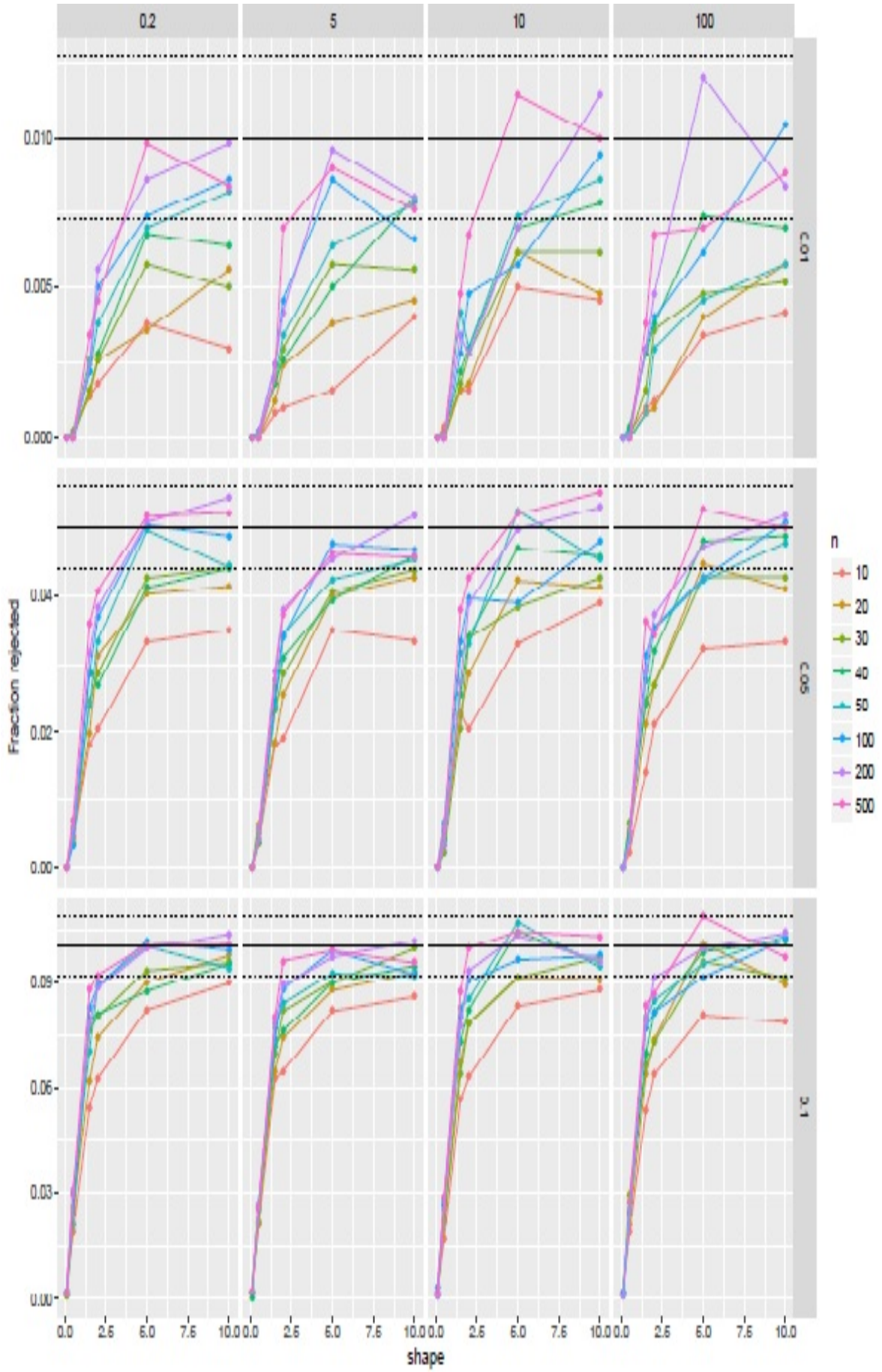
Priloge

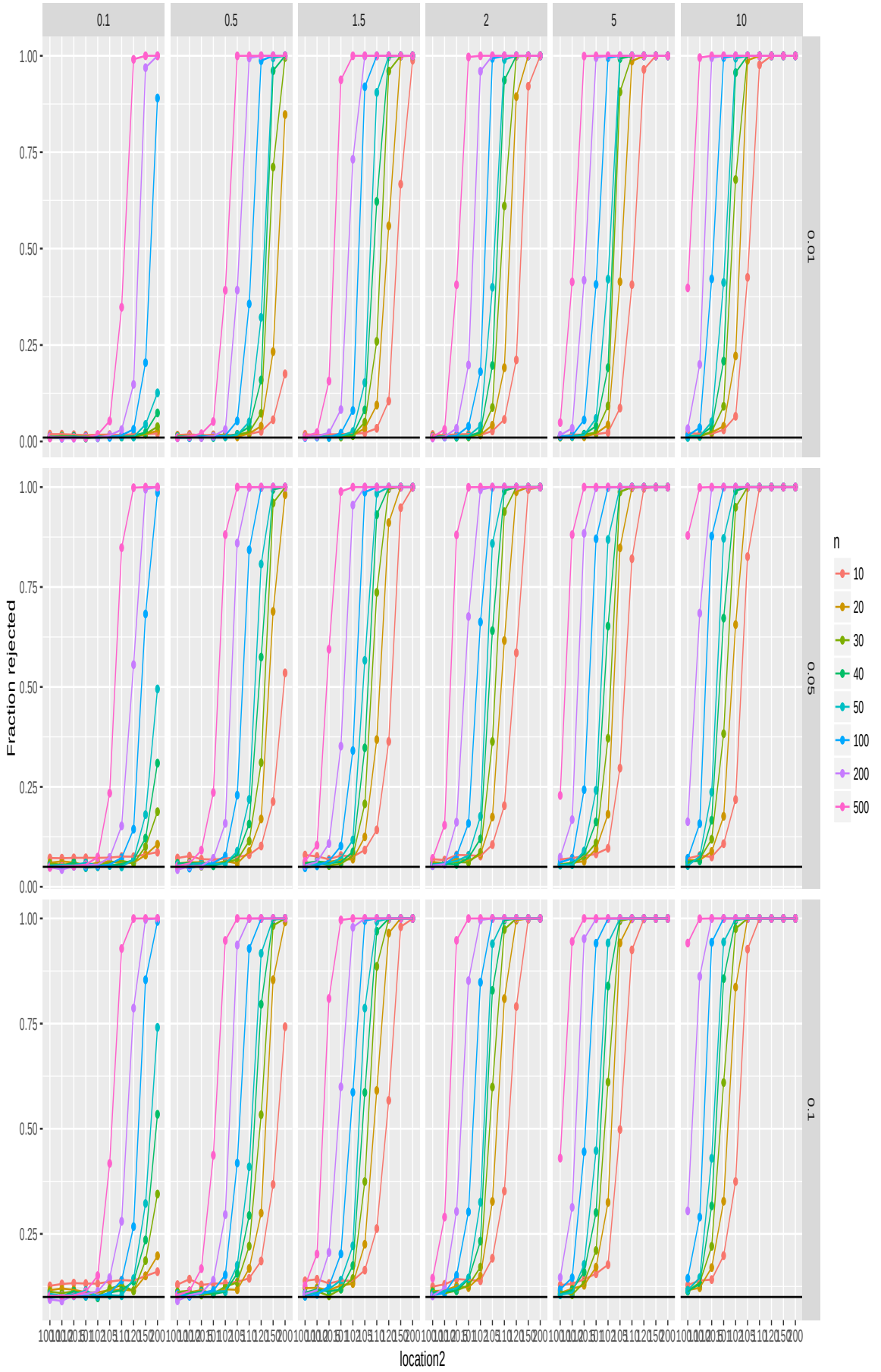
A Dodatni rezultati simulacij

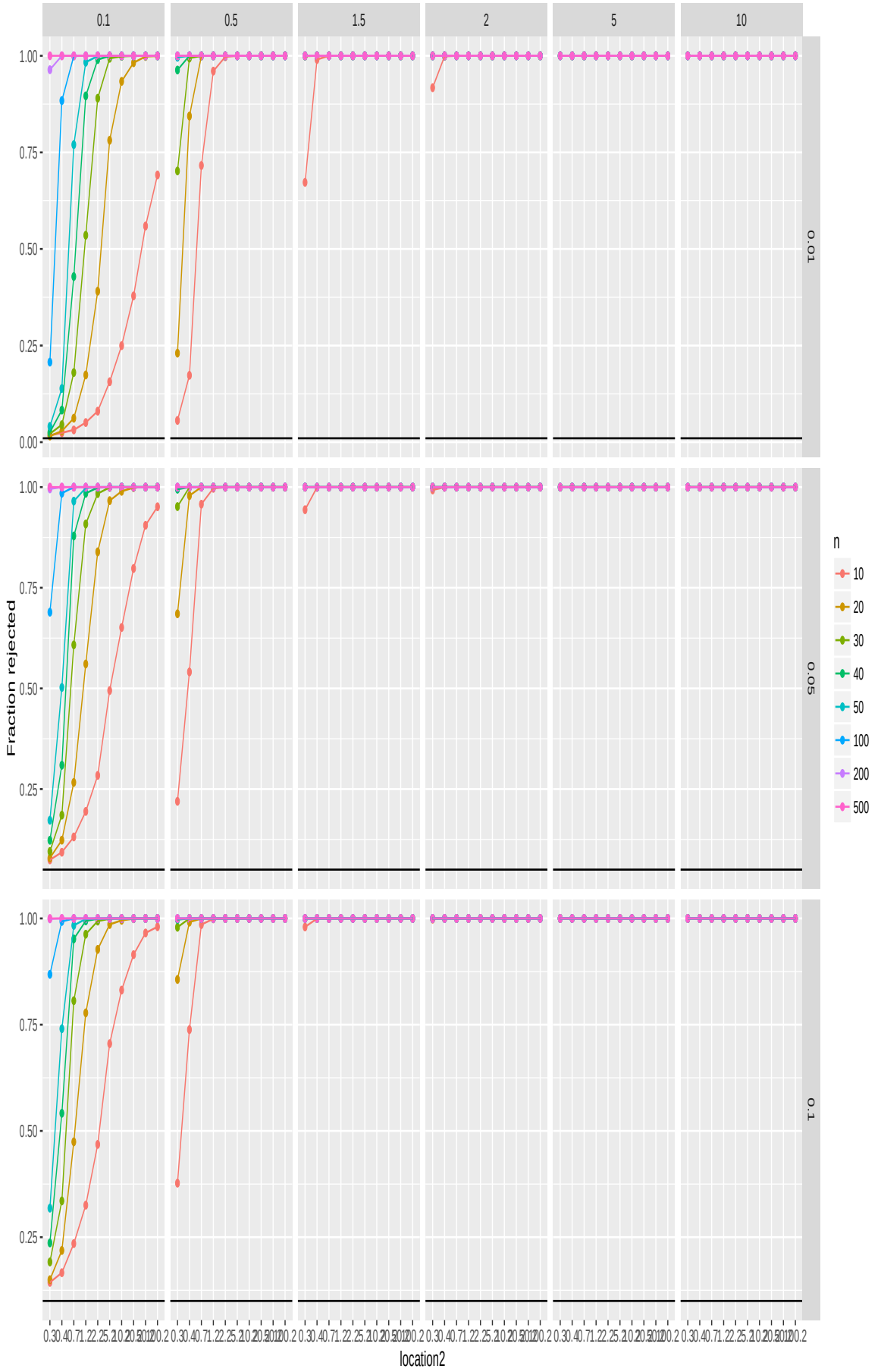

0.2
5
10
100
0.01
0.05
0.1
Fraction rejected
shape
n
10
20
30
40
50
100
200
500

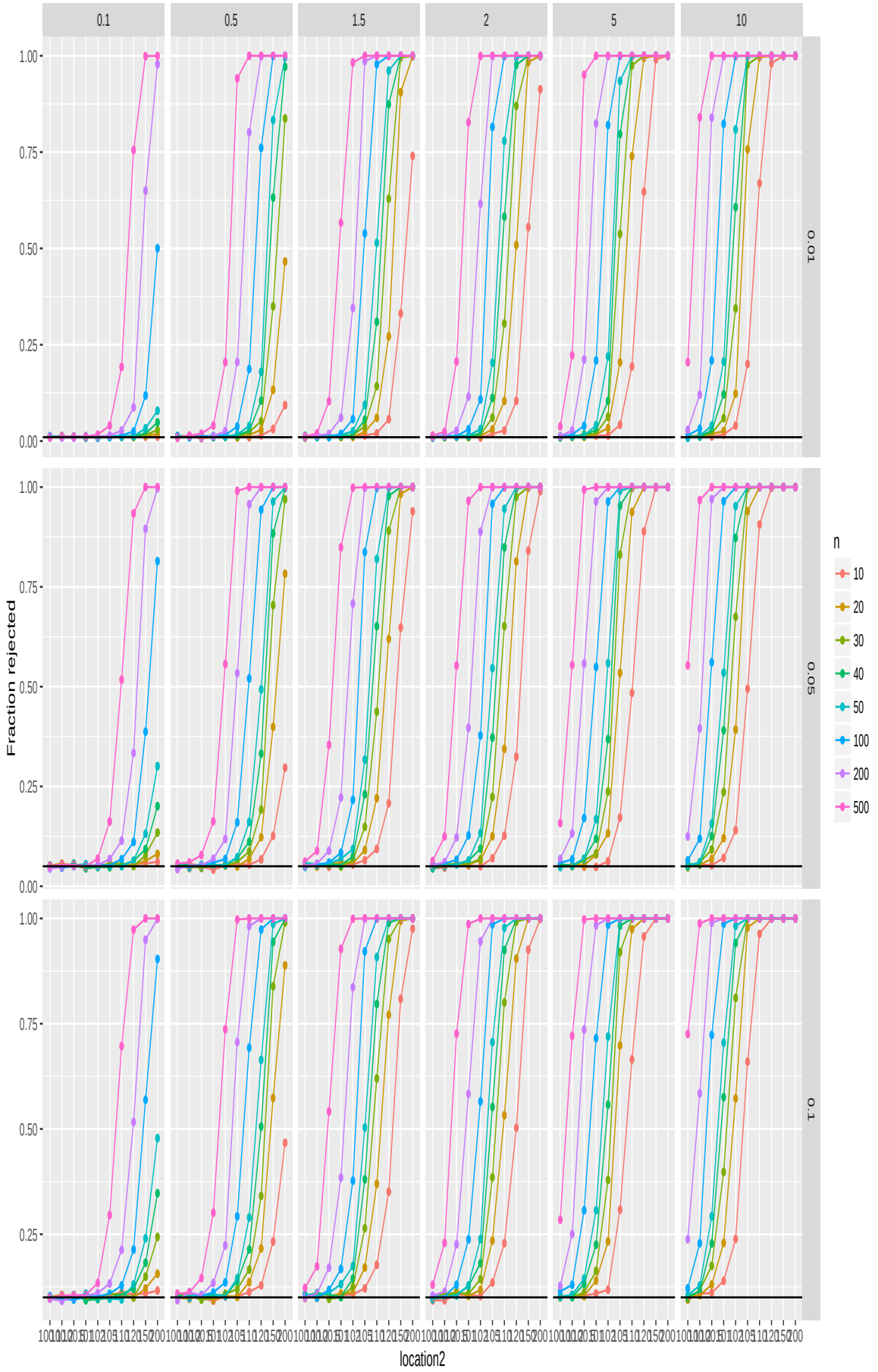

0.2
5
10
100
0.01
0.05
0.1
Fraction rejected
shape
n
10
20
30
40
50
100
200
500

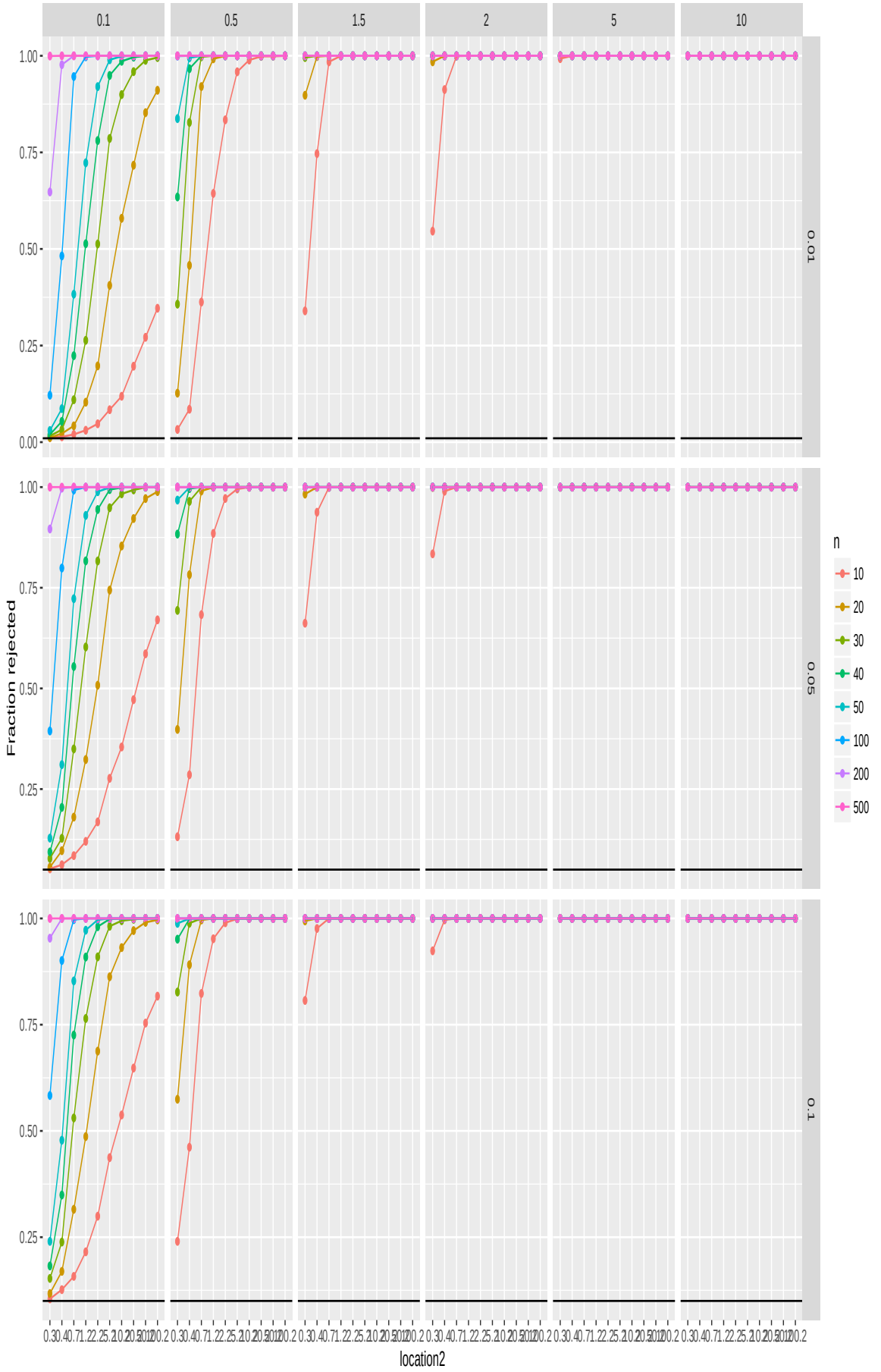

0.2
5
10
100
0.001
0.006
0.1
Fraction rejected
shape
n
10
20
30
40
50
100
200
500

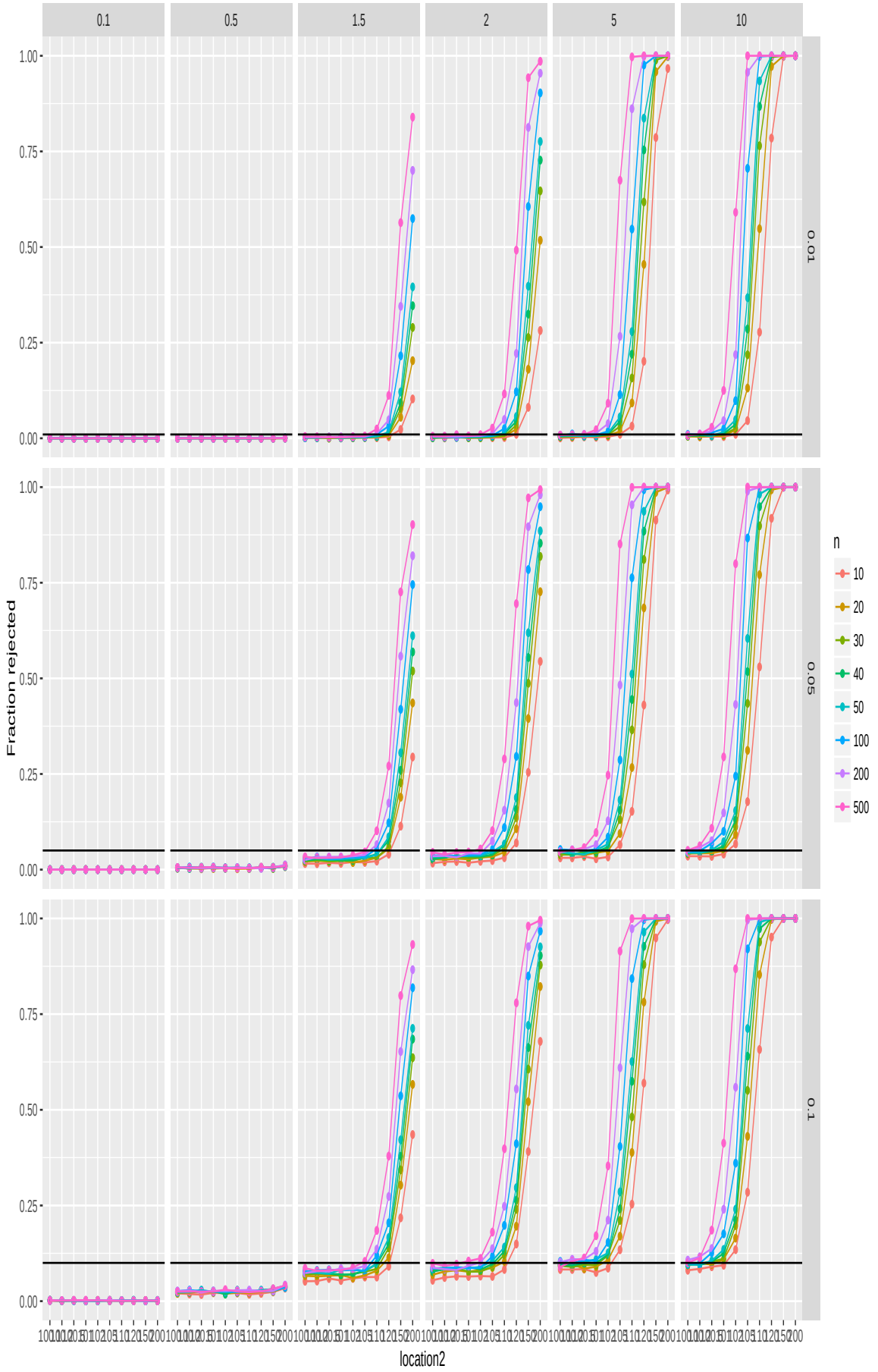

0.2
5
10
100
0.01
0.05
0.1
Fraction rejected
0.010
0.005
0.000
0.04
0.02
0.00
0.09
0.06
0.03
0.00
0.0
2.5
5.0
7.5
10.0
shape
n
10
20
30
40
50
100
200
500

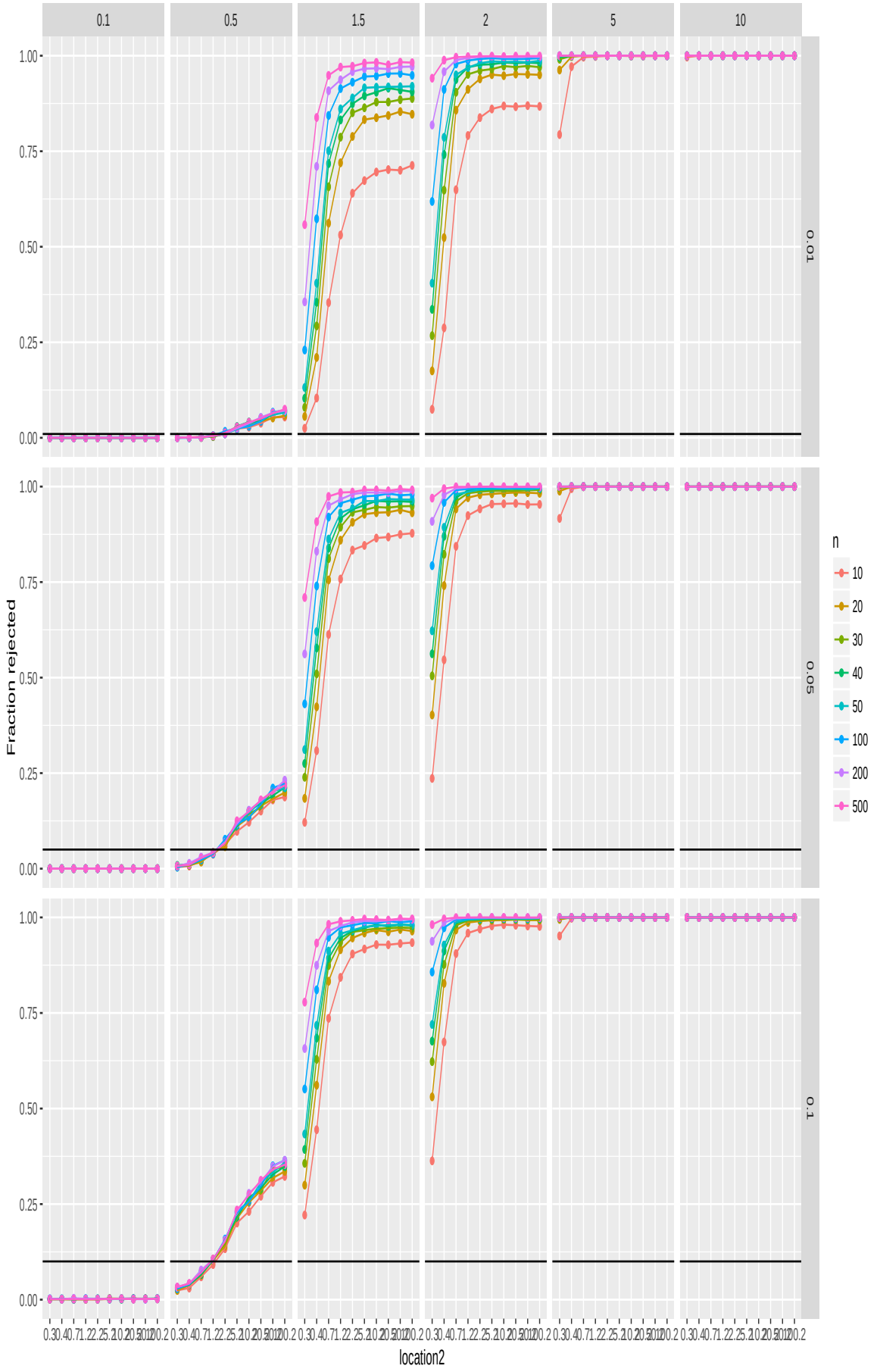

0.1
0.5
1.5
2
5
10
0.01
0.05
0.1
Fraction rejected
location2
n
10
20
30
40
50
100
200
500

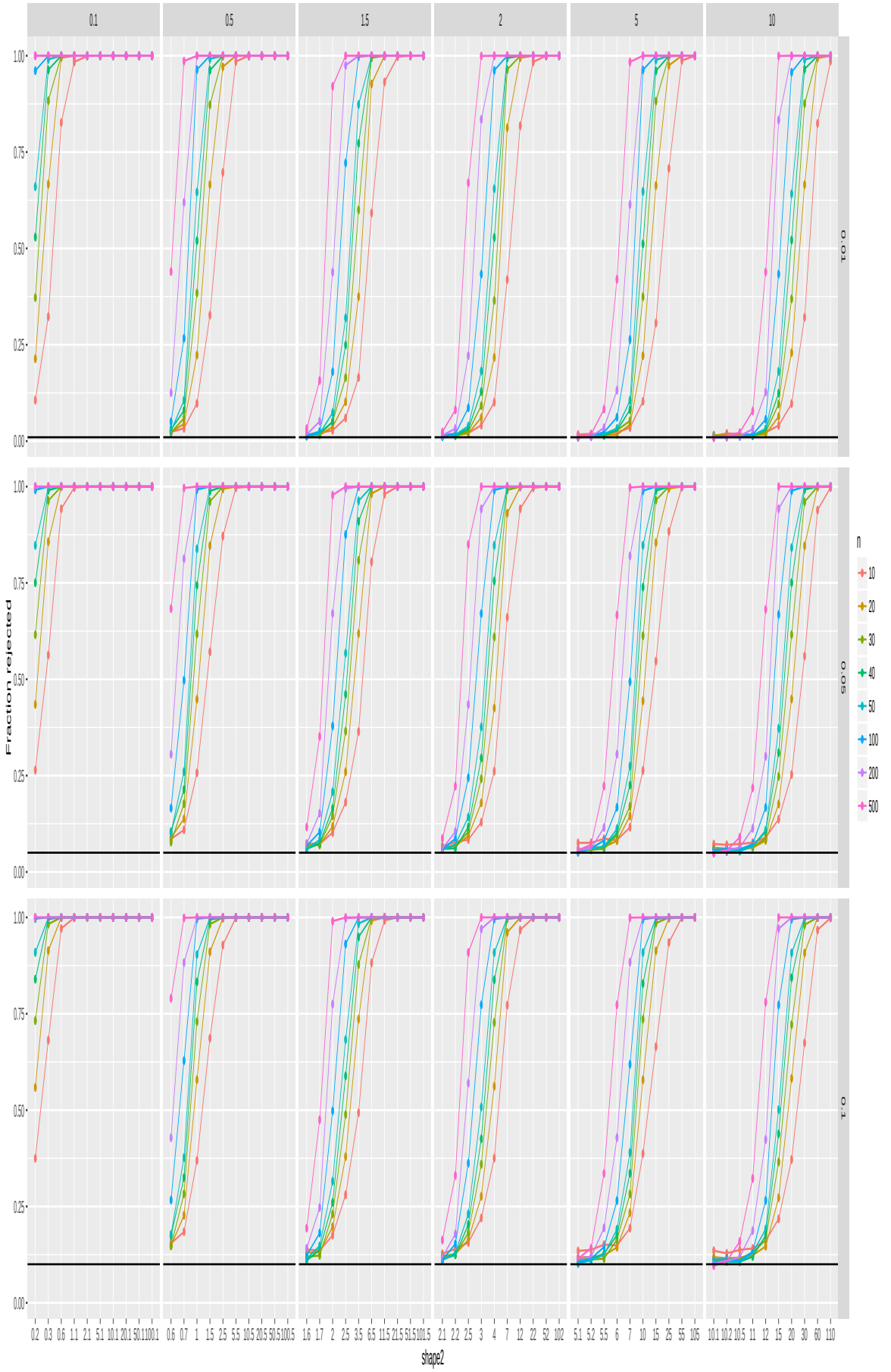

0.1
0.5
1.5
2
5
10
0.01
0.05
0.1
Fraction rejected
1.00
0.75
0.50
0.25
0.00
location2
n
10
20
30
40
50
100
200
500

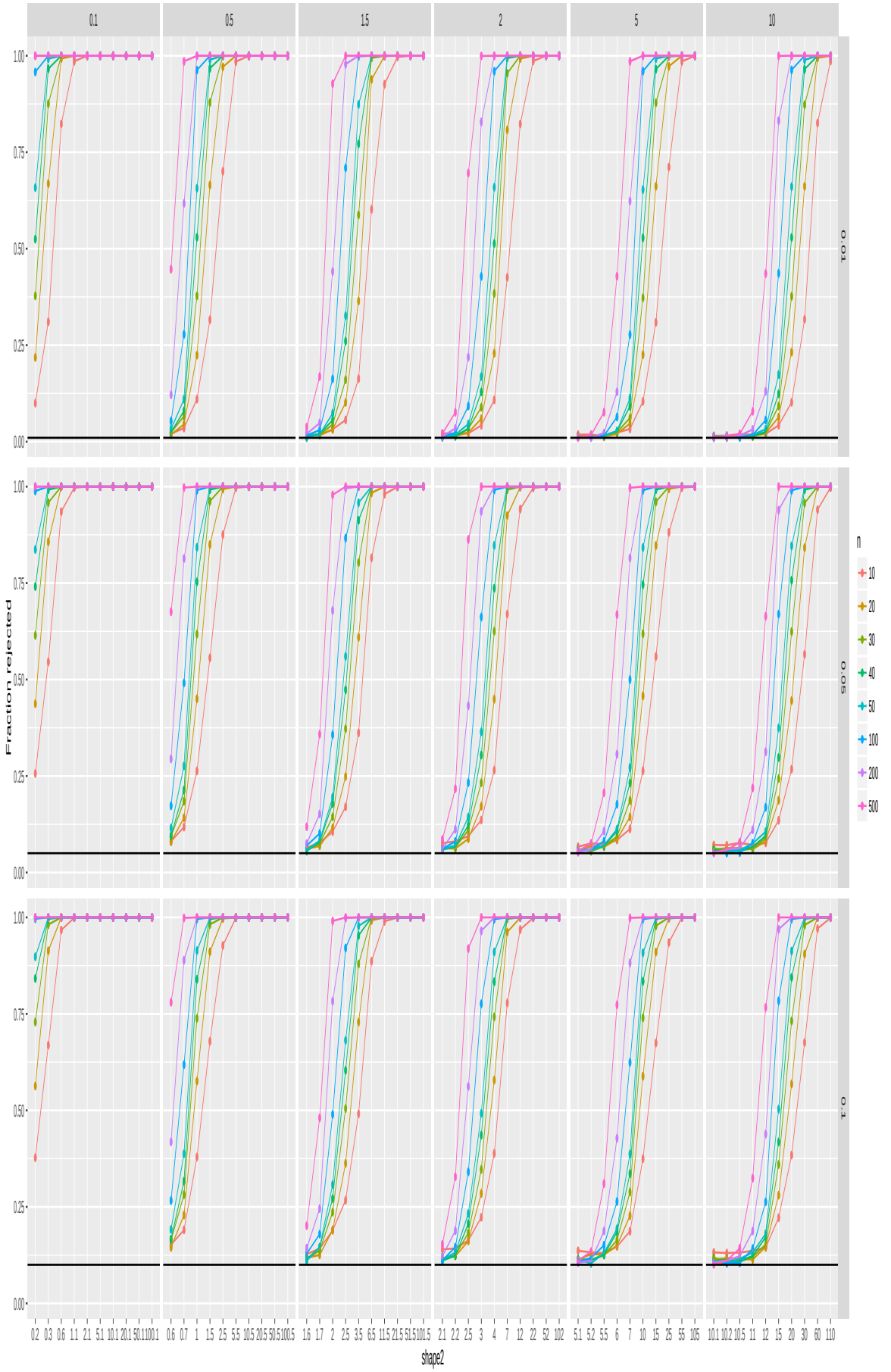

0.1
0.5
1.5
2
5
10
0.01
0.05
0.1
Fraction rejected
location2
n
10
20
30
40
50
100
200
500

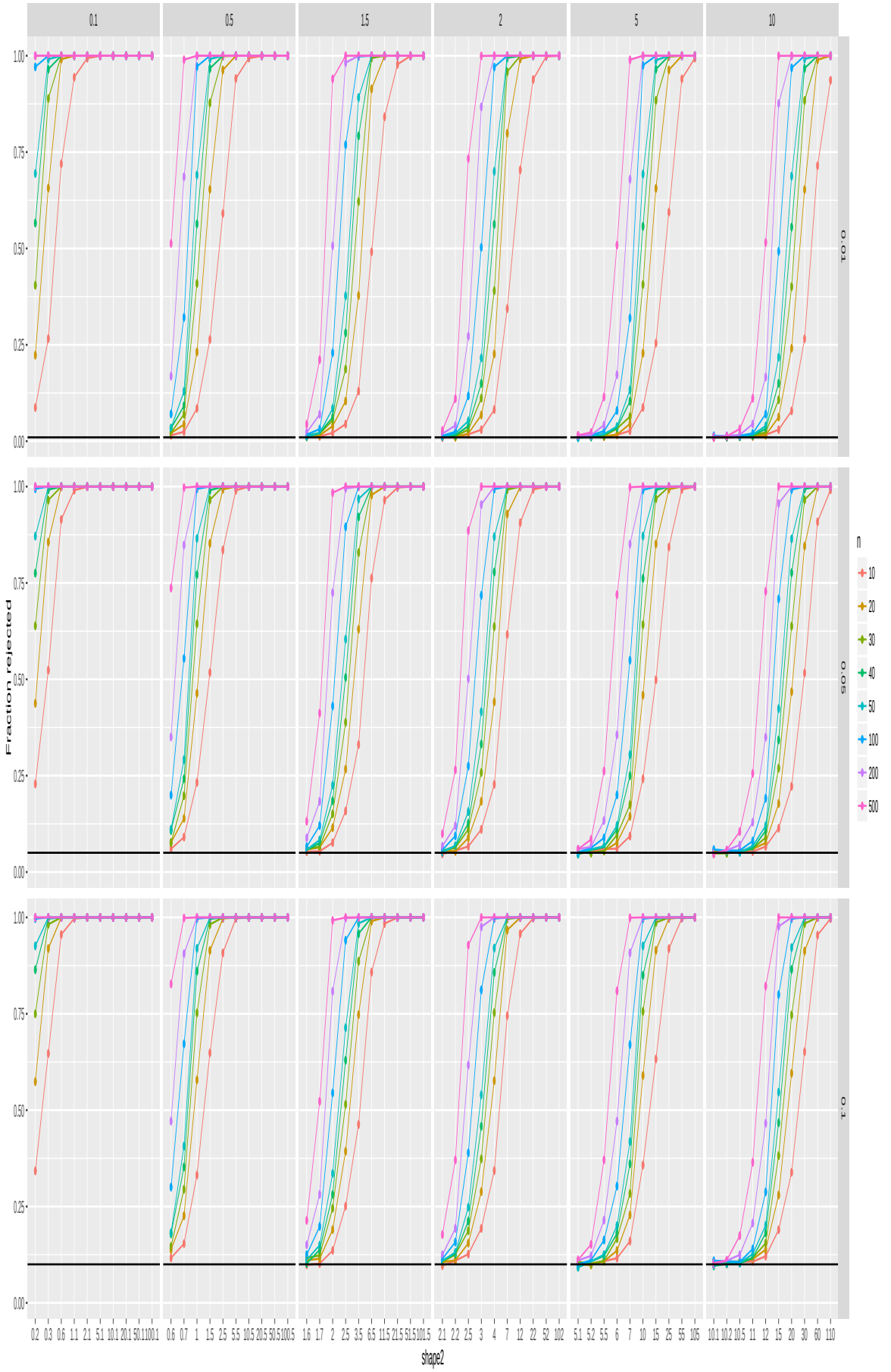

0.1
0.5
1.5
2
5
10
0.01
0.05
0.1
Fraction rejected
1.00
0.75
0.50
0.25
0.00
location2
n
10
20
30
40
50
100
200
500

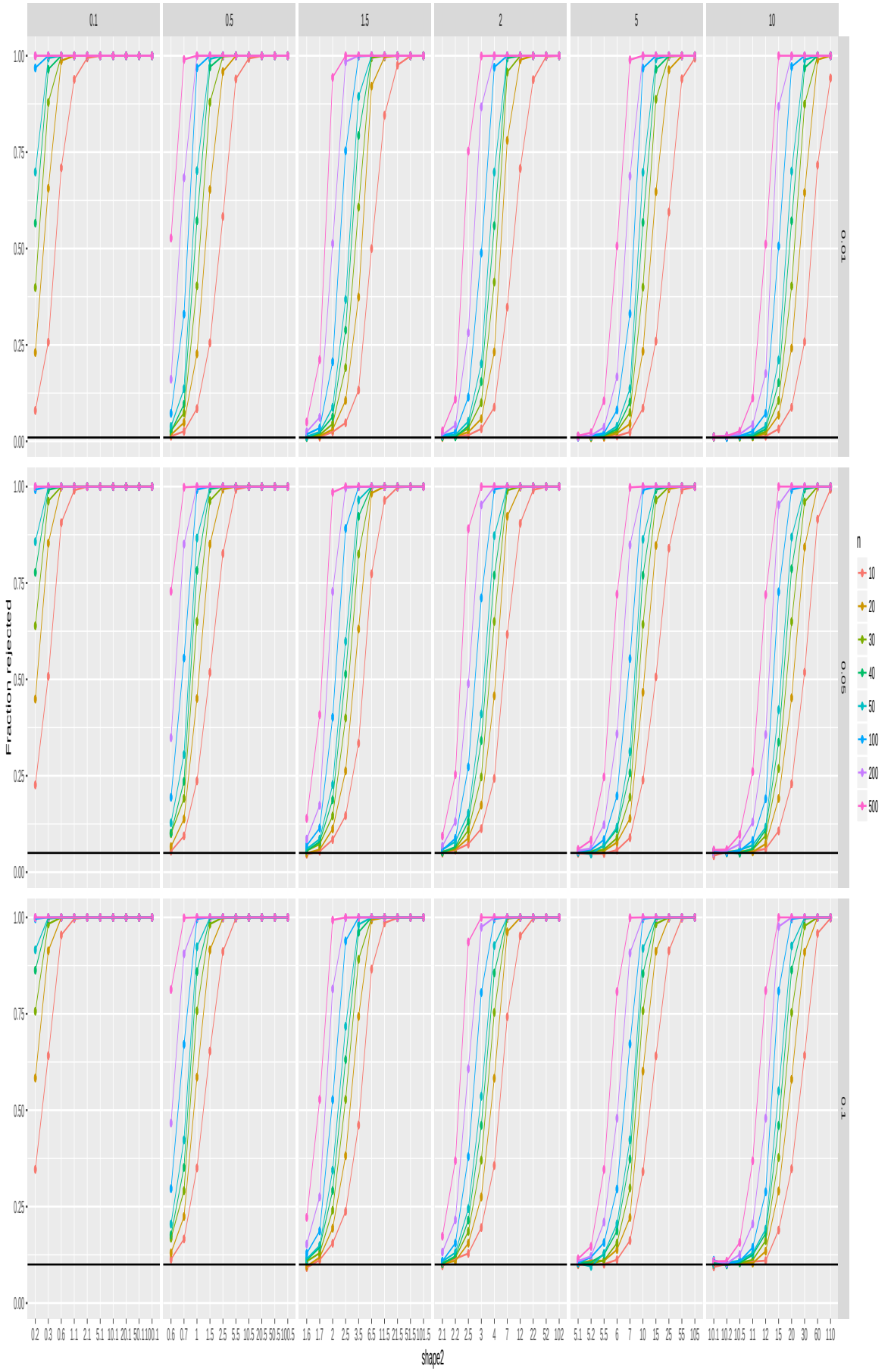

0.1
0.5
1.5
2
5
10
0.01
0.05
0.1
Fraction rejected
location2
n
10
20
30
40
50
100
200
500

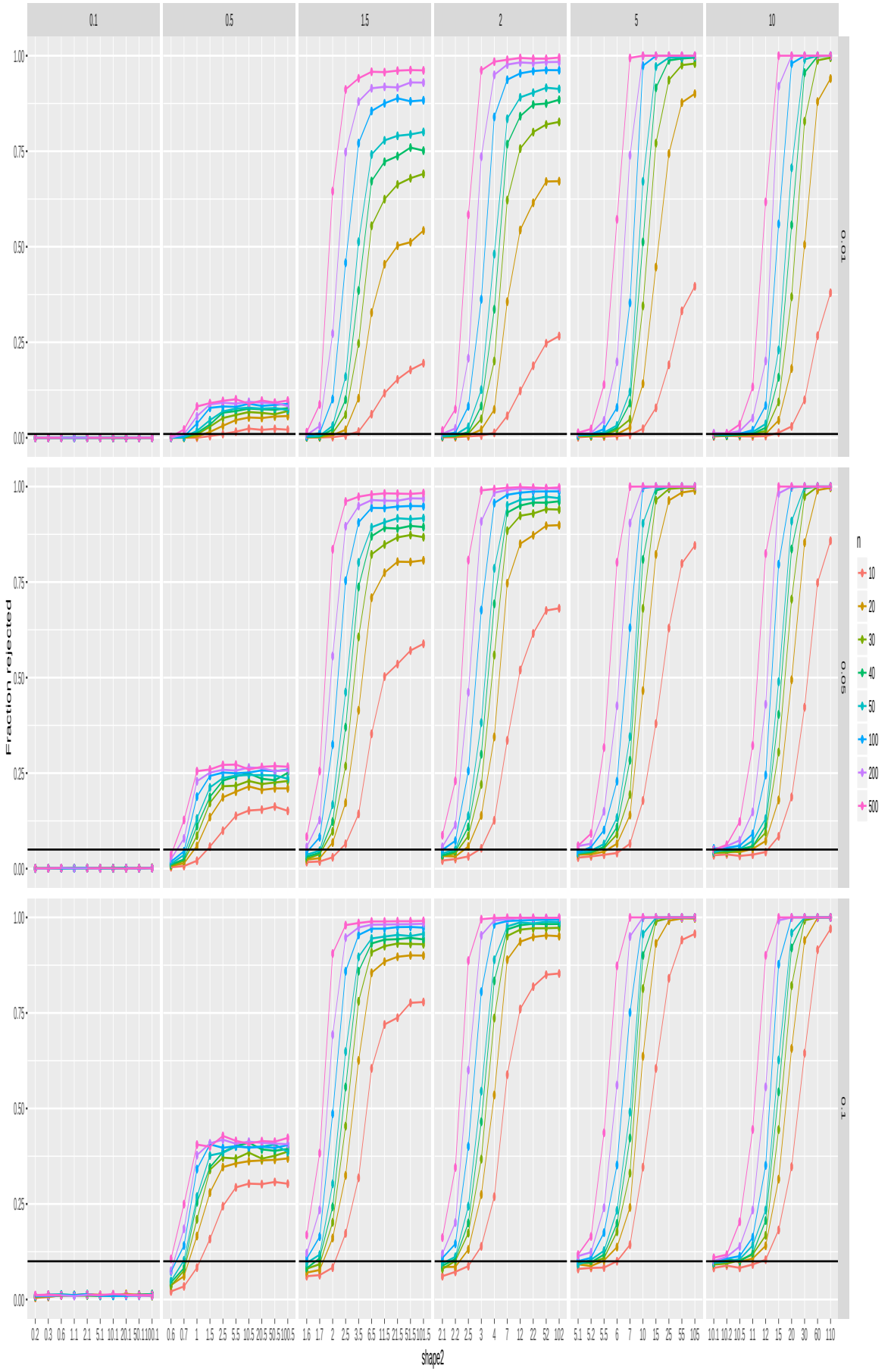

0.1
0.5
1.5
2
5
10
0.01
0.05
0.1
Fraction rejected
location2
n
10
20
30
40
50
100
200
500

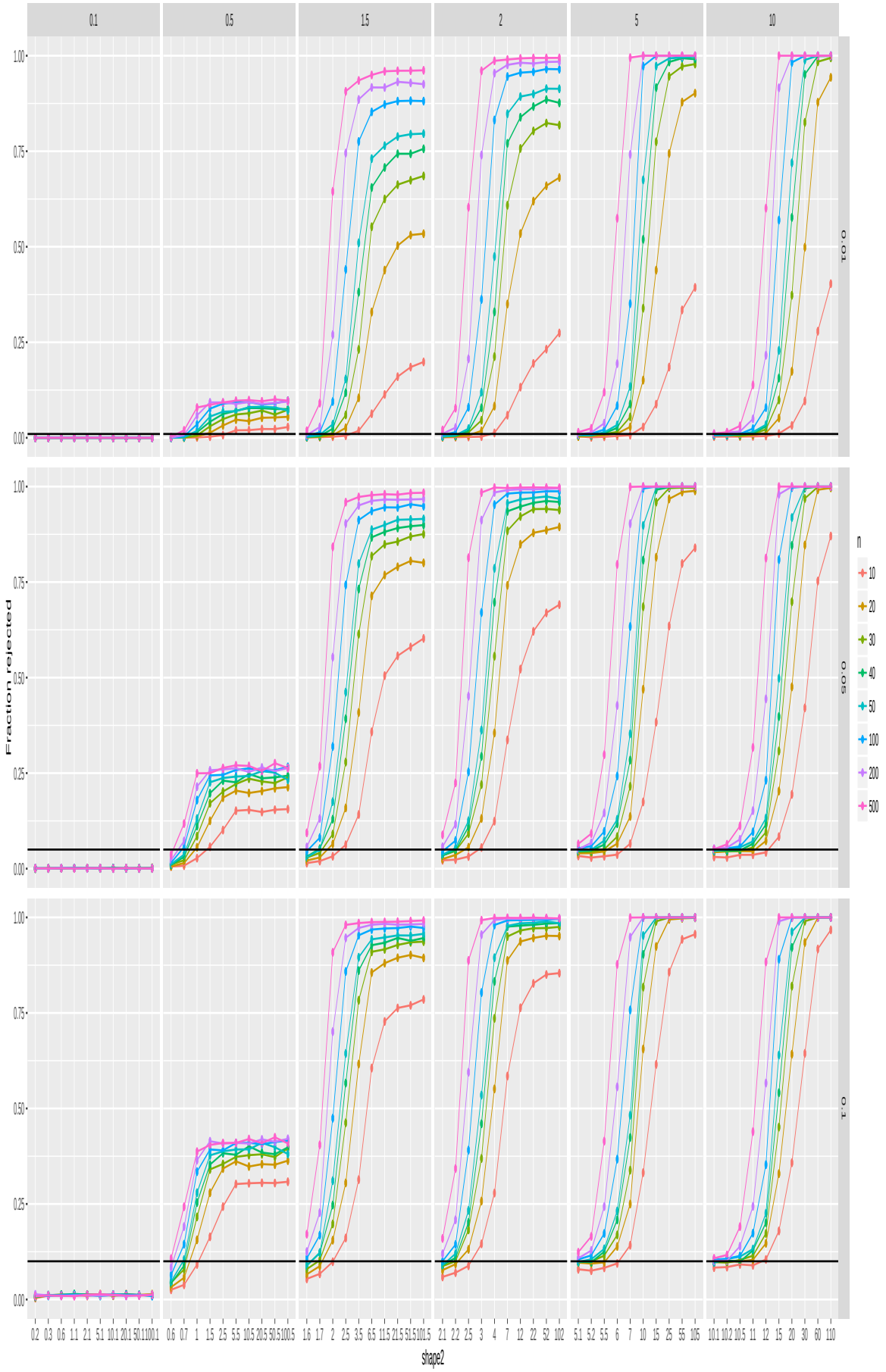

0.1
0.5
1.5
2
5
10
Fraction rejected
0.00
0.25
0.50
0.75
1.00
0.01
0.05
0.1
n
10
20
30
40
50
100
200
500
shape2
0.2 0.3 0.6 1.1 2.1 5.1 10.1 20.1 50.1 100.1
0.6 0.7 1 1.5 2.5 5.5 10.5 20.5 50.5 100.5
1.6 1.7 2 2.5 3.5 6.5 11.5 21.5 51.5 101.5
2.1 2.2 2.5 3 4 7 12 22 52 102
5.1 5.2 5.5 6 7 10 15 25 55 105
10.1 10.2 10.5 11 12 15 20 30 60 110


0.1
0.5
1.5
2
5
10
Fraction rejected
0.00
0.25
0.50
0.75
1.00
shape2
0.2 0.3 0.6 1.1 2.1 5.1 10.1 20.1 50.1 100.1
0.6 0.7 1 1.5 2.5 5.5 10.5 20.5 50.5 100.5
1.6 1.7 2 2.5 3.5 6.5 11.5 21.5 51.5 101.5
2.1 2.2 2.5 3 4 7 12 22 52 102
5.1 5.2 5.5 6 7 10 15 25 55 105
10.1 10.2 10.5 11 12 15 20 30 60 110
n
10
20
30
40
50
100
200
500


0.1
0.5
1.5
2
5
10
Fraction rejected
1.00
0.75
0.50
0.25
0.00
0.01
0.05
0.1
n
10
20
30
40
50
100
200
500
0.2 0.3 0.6 1.1 2.1 5.1 10.1 20.1 50.1 100.1
0.6 0.7 1 1.5 2.5 5.5 10.5 20.5 50.5 100.5
1.6 1.7 2 2.5 3.5 6.5 11.5 21.5 51.5 101.5
2.1 2.2 2.5 3 4 7 12 22 52 102
5.1 5.2 5.5 6 7 10 15 25 55 105
10.1 10.2 10.5 11 12 15 20 30 60 110
shape2


0.1
0.5
1.5
2
5
10
Fraction rejected
0.00
0.25
0.50
0.75
1.00
0.01
0.05
0.1
n
10
20
30
40
50
100
200
500
shape2


0.1
0.5
1.5
2
5
10
Fraction rejected
0.00
0.25
0.50
0.75
1.00
shape2
0.2 0.3 0.6 1.1 2.1 5.1 10.1 20.1 50.1 100.1
0.6 0.7 1 1.5 2.5 5.5 10.5 20.5 50.5 100.5
1.6 1.7 2 2.5 3.5 6.5 11.5 21.5 51.5 101.5
2.1 2.2 2.5 3 4 7 12 22 52 102
5.1 5.2 5.5 6 7 10 15 25 55 105
10.1 10.2 10.5 11 12 15 20 30 60 110
n
10
20
30
40
50
100
200
500
0.01
0.05
0.10


0.1
0.5
1.5
2
5
10
Fraction rejected
0.00
0.25
0.50
0.75
1.00
shape2
0.2 0.3 0.6 1.1 2.1 5.1 10.1 20.1 50.1 100.1
0.6 0.7 1 1.5 2.5 5.5 10.5 20.5 50.5 100.5
1.6 1.7 2 2.5 3.5 6.5 11.5 21.5 51.5 101.5
2.1 2.2 2.5 3 4 7 12 22 52 102
5.1 5.2 5.5 6 7 10 15 25 55 105
10.1 10.2 10.5 11 12 15 20 30 60 110
n
10
20
30
40
50
100
200
500
0.01
0.05
0.1