



M Ú E G Y E T E M 1 7 8 2



Kuruzslás vagy hasznos gyakorlat a szórás (és átlag / átlaggal és) kvartilisekkel való becslése?

Mediya Bawakhan Mrakhan PhD hallgatóval közös cikk alapján

Előadó: Kói Tamás (BME, Semmelweis TMK)

Elméleti (populációs) mennyiségek

Elméleti (populációs) mennyiségek

Metaanalízis távolról

- Transzlációs medicina: orvosi kutatások beépítése a klinikai gyakorlatba
- Metaanalízis: adott témakörben megjelent cikkek statisztikailag megalapozott módon való egyesítése
- Az alapul vett cikkek összegyűjtése szigorú szabályokat követ
- Több cikk együtt nagyobb evidenciát képviselhet
- Feltárja az esetleges ellentmondásokat
- Nemegyszer érdekes mintázatokra is rávilágít

Doing Meta-Analysis in R:

A Hands-on Guide

Table of contents

Welcome!

Preface

About the Authors

Getting Started

1 Introduction

2 Discovering R

Meta-Analysis in R

3 Effect Sizes

4 Pooling Effect Sizes

5 Between-Study Heterogeneity

Welcome!

Welcome to the online version of “Doing Meta-Analysis with R: A Hands-On Guide”.

This book serves as an accessible introduction into how meta-analyses can be conducted in *R*. Essential steps for meta-analysis are covered, including pooling of outcome measures, forest plots, heterogeneity diagnostics, subgroup analyses, meta-regression, methods to control for publication bias, risk of bias assessments and plotting tools.

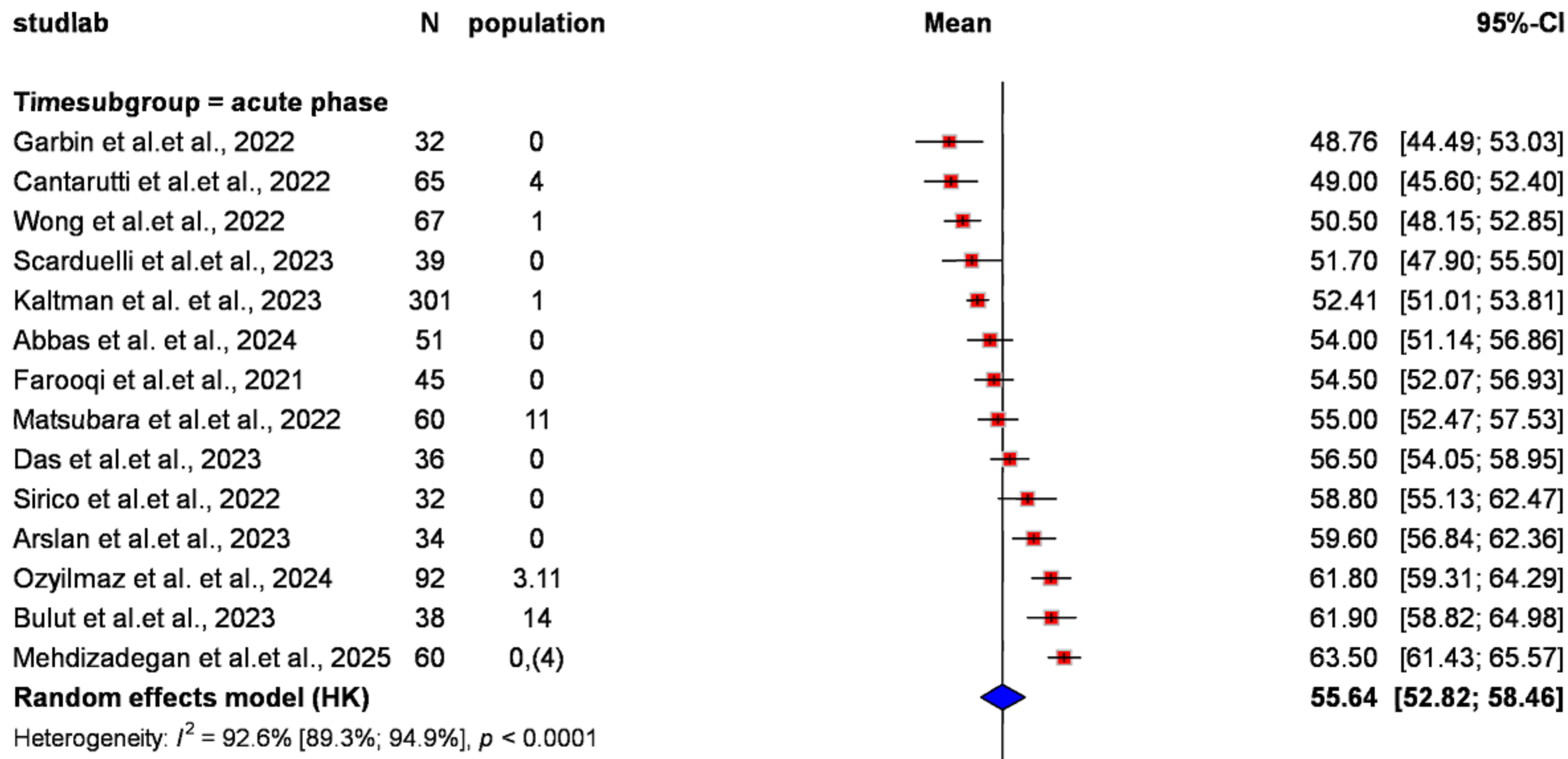
Doing Meta-Analysis with R

A Hands-On Guide



Mathias Harrer
Pim Cuijpers
Toshi Furukawa
David Ebert

First author	Year of	Timesubgro	Num	Mean	Median	SD	SE	Confi	Confid	Lower	Upper
Arslan et al.	2023	acute phase	34	59.6		8.2					
Arslan et al.	2023	mid-term	17	66.5		6.3					
Awasthi et al.	2022	acute phase	40		45					36	50
Awasthi et al.	2022	mid-term	34		60					55	65
Aziz et al.	2022	acute phase	41		51						
Aziz et al.	2022	long-term	38		70						
Bulut et al.	2023	acute phase	38	61.90		9.7					
Bulut et al.	2023	mid-term	34	71.20		4.2					
Cantarutti et al.	2022	acute phase	65	49.00		14					
Cantarutti et al.	2022	short-term	57	62.00		4					
Das et al.	2023	acute phase	36	56.50		7.5					
Das et al.	2023	mid-term	17	60.40		3.5					
Farooqi et al.	2021	acute phase	45	54.50		8.3					
Farooqi et al.	2021	short-term	31	60.50		4.40					
Farooqi et al.	2021	mid-term	24	63.80		4					
Garbin et al.	2022	mid-term	32	63.3		4.50					
Garbin et al.	2022	acute phase	32	48.763919		12.32					
Kaltman et al.	2023	long-term	5	64.97		1.27					
Kaminska et al.	2023	acute phase	172		63					54	68
Kaminska et al.	2023	mid-term	88		63					62	65
Krasic et al.	2023	acute phase	67		58					52	68
Krasic et al.	2023	mid-term	63	66.40		4.1					
Matsubara et al.	2022	acute phase	60	55.00		10					
Matsubara et al.	2022	short-term	25	64.00		5					
Mehdizadegan et	2025	acute phase	60	63.50		8.196					
Mehdizadegan et	2025	short-term	55	70.12		4.301					



„Random-effect” metaanalízis

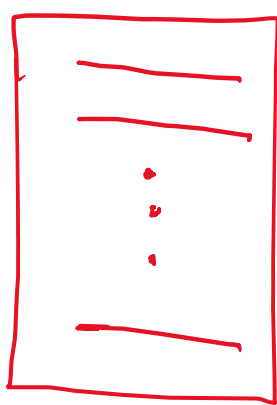
$$\hat{\theta}_k = \boxed{\mu} + \zeta_k + \epsilon_k$$

PARAMÉTER

$$\zeta_k \sim N(0, \boxed{\gamma}), \quad \epsilon_k \sim N(0, \sigma_k^2)$$

PARAMÉTER

STUDY 1 : $M + \zeta_1$ ELMÉLETI PAR



$$\Rightarrow \begin{pmatrix} \hat{\theta}_1 & \hat{\sigma}_1 \end{pmatrix}$$

STUDY 2 : $M + \zeta_2$ EUR. PAR



$$\Rightarrow \begin{pmatrix} \hat{\theta}_2 & \hat{\sigma}_2 \end{pmatrix}$$

...

ALGORITMUS EREDMÉNYE : $\hat{\mu}, SE_{\hat{\mu}}, \hat{\gamma}$

Korábbi cikkek

Szcenáriók

$$S1 = \{Q_{\min}, Q_2, Q_{\max}, n\}$$

$$S2 = \{Q_1, Q_2, Q_3, n\}$$

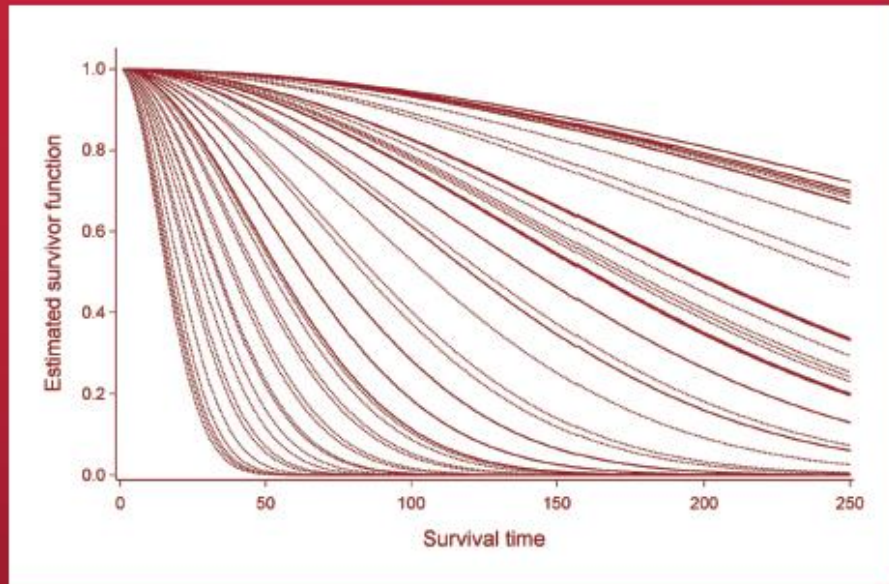
$$S3 = \{Q_{\min}, Q_1, Q_2, Q_3, Q_{\max}, n\}$$

Study	Estimated statistics	Covered Scenarios	Comment
Hozo et al. 2005 [7]	\bar{X}, S^*	S1	Explicit formulas
Bland 2014 [2]	\bar{X}, S^*	S1, S2, S3	Generalization of Hozo et al. [7]
Wan et al. 2014 [18]	\bar{X}, S^*	S1, S2, S3	Sample size is taken into account, explicit formulas
Kwan et al. 2015 [8]	\bar{X}, S^*	S1, S2, S3	Bayesian, min-max range and IQR can be used with modifications
Luo et al. 2016 [9]	\bar{X}	S1, S2, S3	Improved mean estimations using weight optimization
Shi et al. 2020 [15]	\bar{X}, S^*	S1, S2, S3	Connected closely to Luo et al. [9] improved SD estimations using weight optimization
McGrath et al. 2020 [12]	\bar{X}, S^*	S1, S2, S3	Introduced QE and a Box Cox transformation based method
Cai et al. 2021 [4]	\bar{X}, S^*	S1, S2, S3	Method based on maximum likelihood and Box Cox transformation, provides two algorithms to estimate the power parameter in BC
Balakrishnan et al. 2022 [1]	\bar{X}, S^*	S1, S2, S3	Arbitrary ordered statistics can be used, e.g., 10%- and 90%-percentiles can be incorporated in the estimation.
McGrath et al. 2023 [10]	SE of the est. mean	S1, S2, S3	Using parametric bootstrap it improve standard error estimates of estimated means

Általános kép

Általános kép

Modelling Survival Data in Medical Research Third Edition



David Collett

Delta formula

$$\text{var} \{g(X)\} \approx \left\{ \frac{dg(X)}{dX} \right\}^2 \text{var}(X).$$

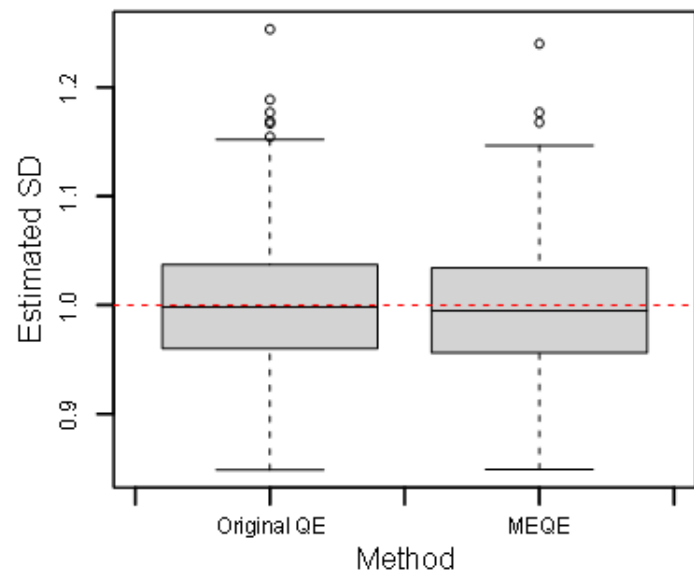
R-ben a default beállítás szerint a t-idős túlélés valószínűség konfidencia intervalluma log transzformációval készül

$$\text{var} \left\{ \log \hat{S}(t) \right\} \approx \frac{1}{[\hat{S}(t)]^2} \text{var} \left\{ \hat{S}(t) \right\}$$

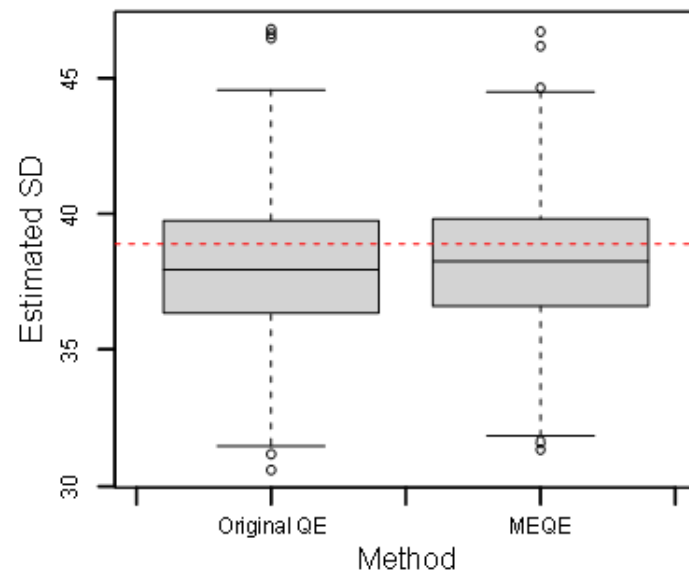
Átlaggal kiegészített QE, S2 scenárió

- Normál eloszlás család, $\theta_N = (\mu, \sigma)$
- Lognormális eloszlás $\theta_{LN} = (\mu, \sigma)$
- Gamma eloszlás család, $\theta_G = (\alpha, \lambda)$, α alak paraméter, λ ráta paraméter
- Weibull eloszlás család, $\theta_W = (k, \lambda)$, k alak paraméter, λ skála paraméter
- Mind a négy paramétercsaládra megkeressük melyik θ paraméter minimalizálja a lenti négyzetes célfüggvényt
- A négy közül kiválasztjuk azt amihez tartozó célfüggvény a legkisebb
- A kapott optimális család optimális paramétere alapján számolunk elméleti várható értéket és szórást
- Egyszerűsége miatt könnyű volt bevenni az átlagot: estmeansd R csomag viszonylag egyszerű módosításával oldottuk meg

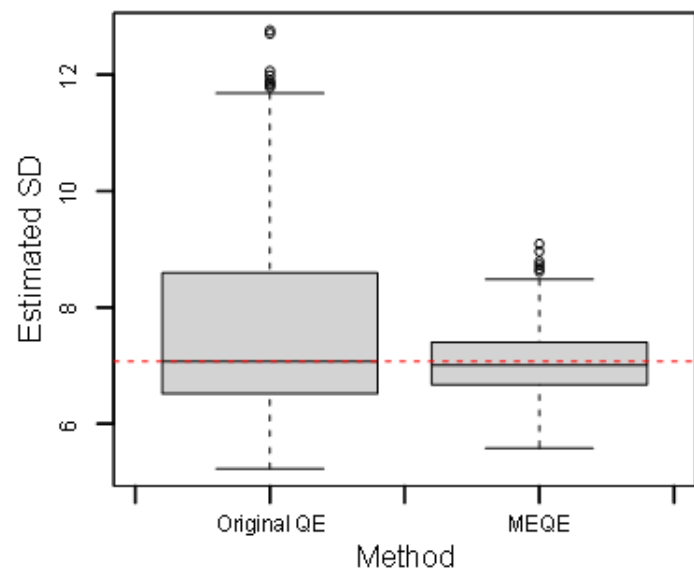
$$S_2(\theta) = (F_\theta^{-1}(0.25) - Q_1)^2 + (F_\theta^{-1}(0.5) - Q_2)^2 + (F_\theta^{-1}(0.75) - Q_3)^2 + \left(\mu_\theta^{TH} - \bar{X} \right)^2$$



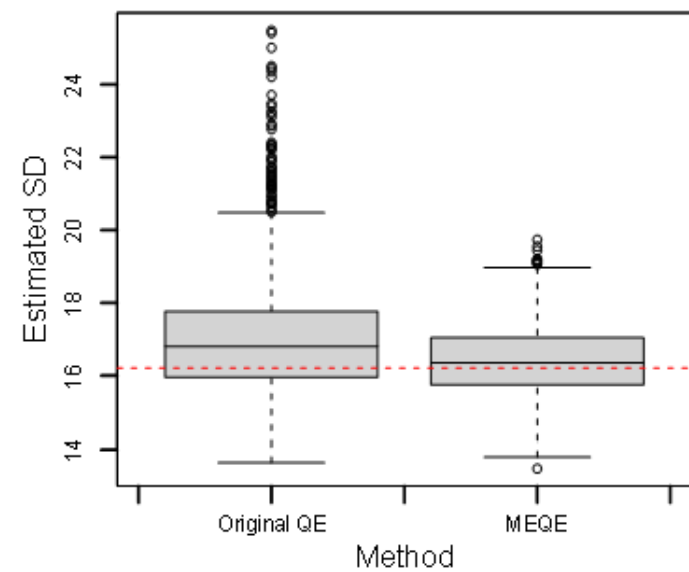
Normal (5, 1)



Log-normal (5, 0.25)



Gamma (2, 5)



Weibull (2, 35)

ARE (average relative error) kritériumok

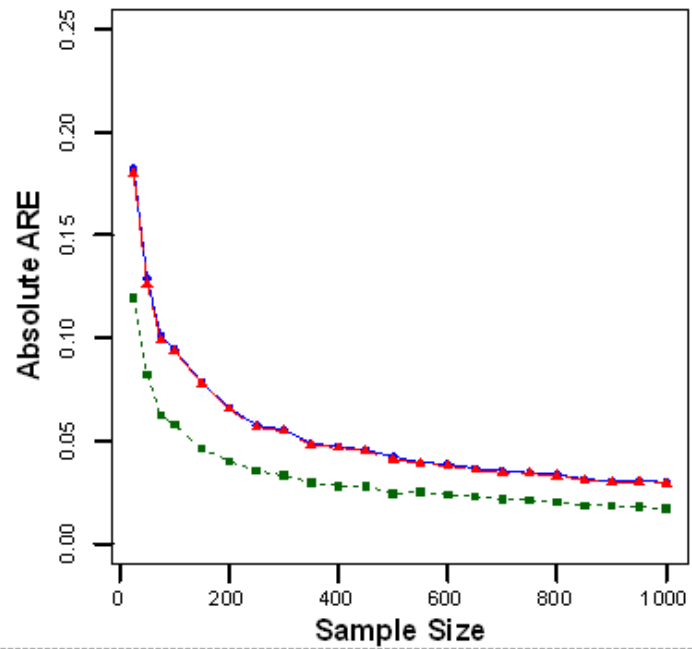
- *1000-szer szimuláltunk n elemű mintát*
- \hat{S}_i az i . szimuláció esetén adódó becslés
- S_i^* az i . szimulációban a mintából számolt klasszikus szórás
- s^{TH} a háttérváltozó elméleti szórása

$$\text{ARE}_{\{est-sample\}} = \frac{1}{1000} \sum_{i=1}^{1000} \frac{\hat{S}_i - S_i^*}{S_i^*} \quad \text{ARE}_{\{est-sample\}}^{abs} = \frac{1}{1000} \sum_{i=1}^{1000} \left| \frac{\hat{S}_i - S_i^*}{S_i^*} \right|$$

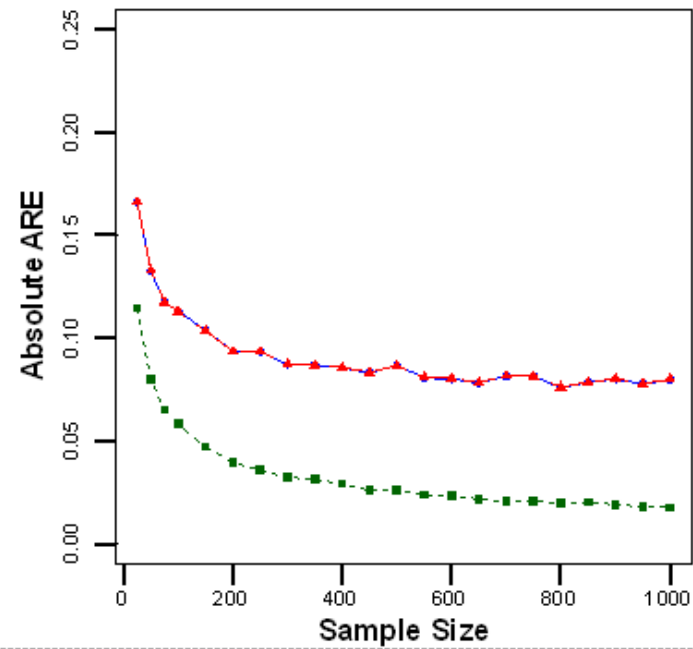
$$\text{ARE}_{\{est-theo\}} = \frac{1}{1000} \sum_{i=1}^{1000} \frac{\hat{S}_i - s^{TH}}{s^{TH}} \quad \text{ARE}_{\{est-theo\}}^{abs} = \frac{1}{1000} \sum_{i=1}^{1000} \left| \frac{\hat{S}_i - s^{TH}}{s^{TH}} \right|$$

A második sor két kifejezés S_i^* -ra vagyis a mintából számolt klasszikus (korrigálatlan) szórásra is felírható

Scenario S2



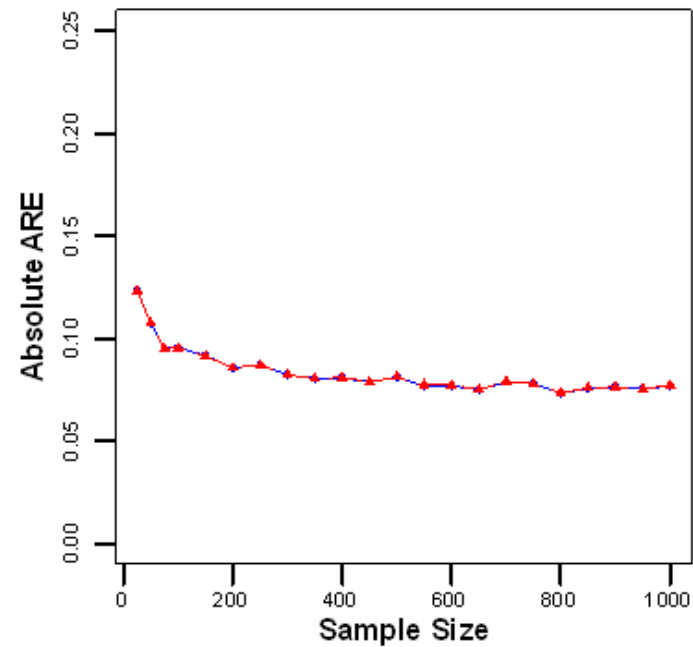
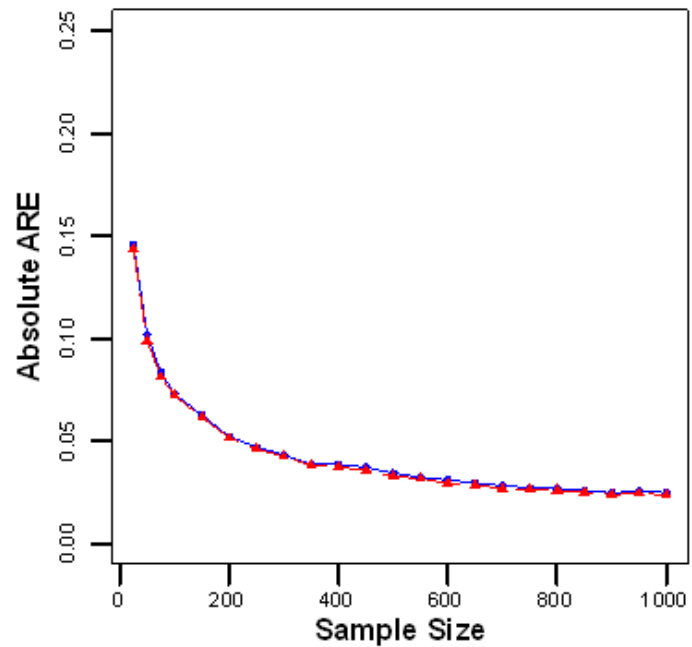
Scenario S3



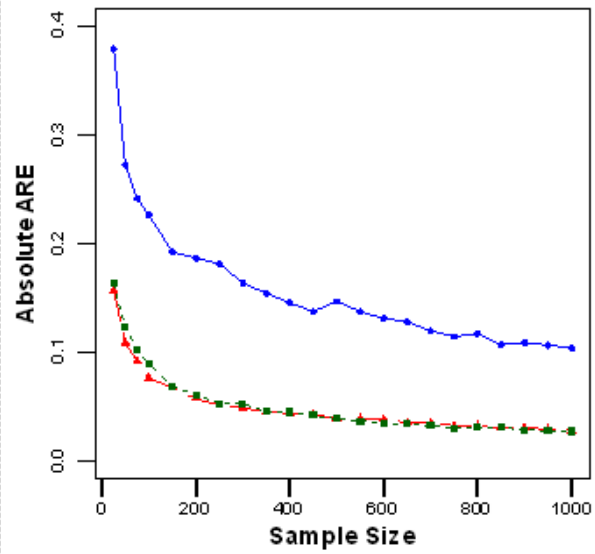
Viszonyítás: elméleti szórás

A háttérváltozó (5,1) paraméterű normális eloszlást követ

Viszonyítás: mintából számolt szórás

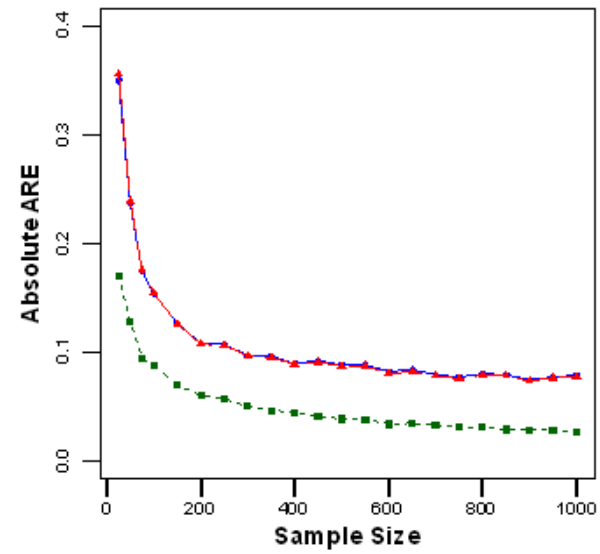


Scenario S2

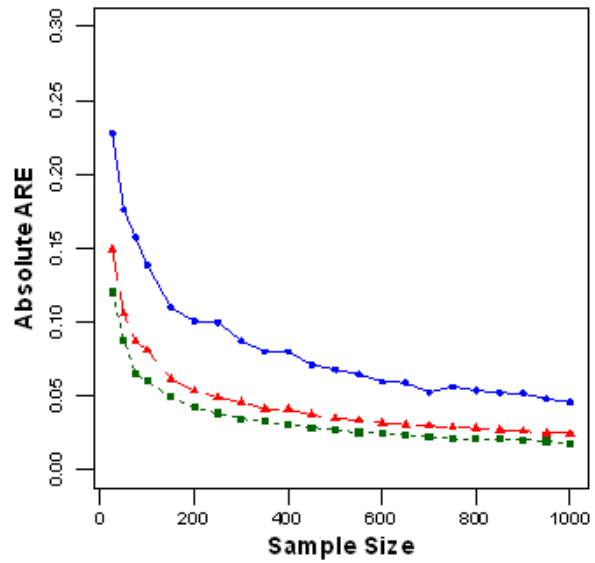


Gamma (2, 5)

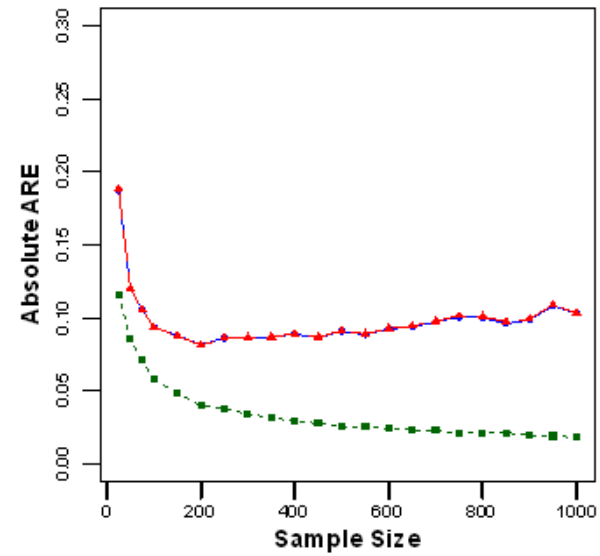
Scenario S3



Gamma (2, 5)



Weibull (2, 35)



Weibull (2, 35)

Viszonyítás: elméleti szórás

A Note on the Asymptotic Distribution of Sample Quantiles

By A. M. WALKER

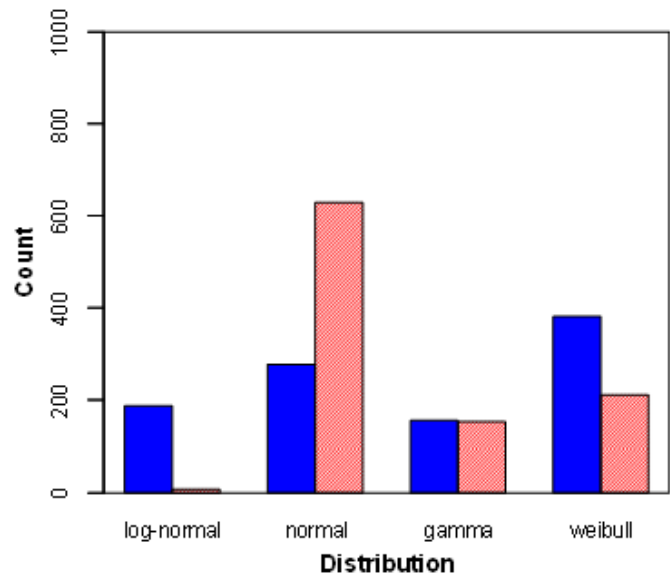
Statistical Laboratory, University of Cambridge†

[Received April 1967. Revised September 1967]

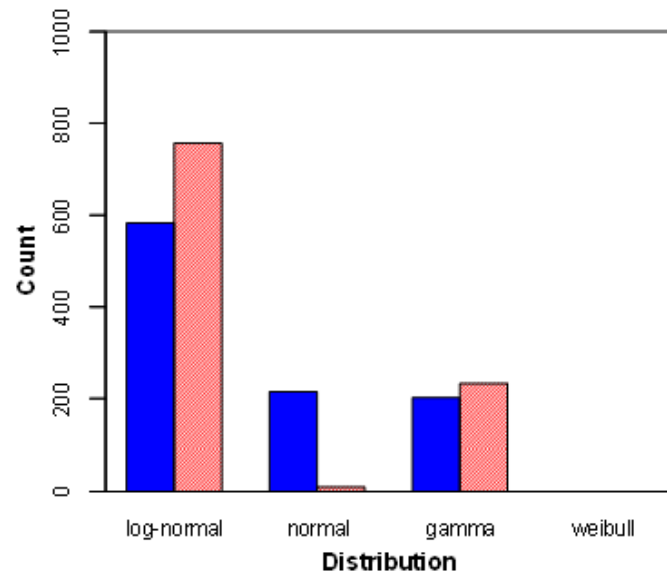
Szép esetben $0 < p < 1$ -re nagy n esetén az empirikus p -kvantilis szórása körülbelül

$$\text{SE}(\hat{\xi}_p) \approx \sqrt{\frac{p(1-p)}{nf(\xi_p)^2}}.$$

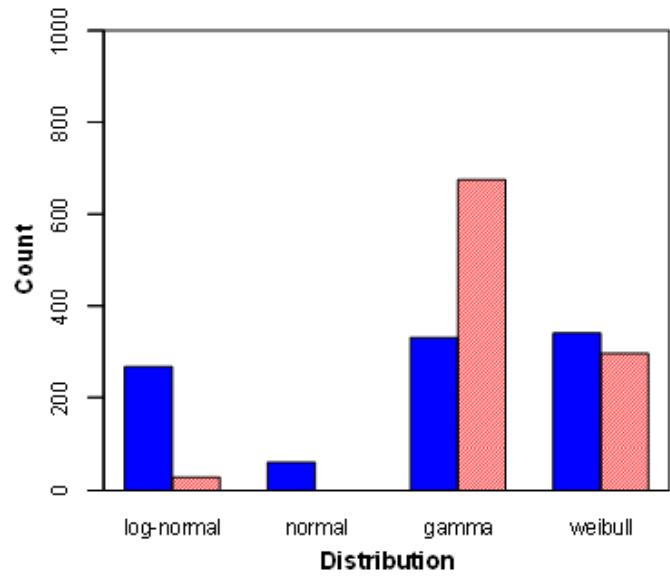
Így ha jó az eloszlásbecslés akkor lehetőség van meta-analizálni a mediánt és a kvartiliseket is (utóbbi tesztelése folyamatban)



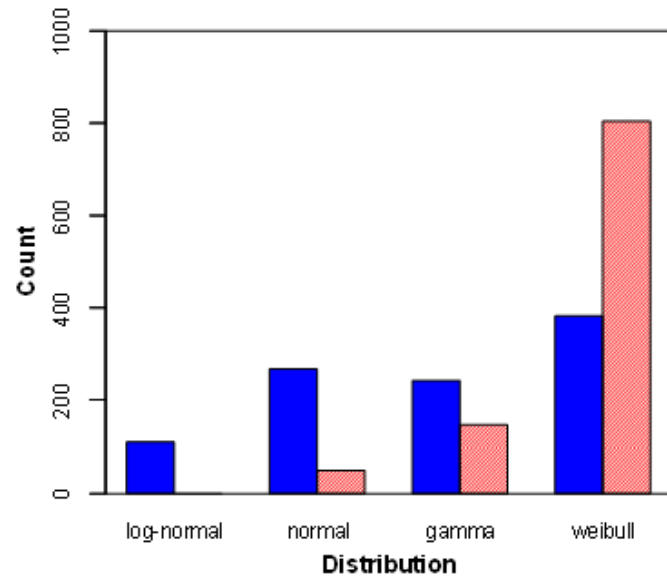
Normal (5, 1)



Log-normal (5, 0.25)



Gamma (2, 5)



Weibull (2, 35)

Az eloszlás detekciót külön is vizsgáltuk

- Igazából nem a minimum hanem a maximum rontja a teljesítményt
- Az egyik lehetséges magyarázatot az extrémérték elmélet szolgáltatja
- Valós adatokon az outlier-ek további problémát okozhatnak

APPLICATIONS
OF MATHEMATICS
STOCHASTIC
MODELLING
AND APPLIED
PROBABILITY

33

Paul Embrechts
Claudia Klüppelberg
Thomas Mikosch

Modelling Extremal Events

for Insurance and Finance



 Springer

FAE valószínűségi változók összegének jobb megértése

FAE valószínűségi változók összegének jobb megértése

X és Y azonos típusú vv-k ha léteznek $b > 0$ és $a \in \mathbb{R}$, hogy

$$X \stackrel{d}{=} bY + a$$

Theorem A1.5 (Convergence to types theorem)

Let A, B, A_1, A_2, \dots be rvs and $b_n > 0, \beta_n > 0$ and $a_n, \alpha_n \in \mathbb{R}$ be constants.

Suppose that

$$b_n^{-1} (A_n - a_n) \xrightarrow{d} A.$$

Then the relation

$$\beta_n^{-1} (A_n - \alpha_n) \xrightarrow{d} B \tag{A.2}$$

holds if and only if

$$\lim_{n \rightarrow \infty} b_n / \beta_n = b \in [0, \infty), \quad \lim_{n \rightarrow \infty} (a_n - \alpha_n) / \beta_n = a \in \mathbb{R}. \tag{A.3}$$

If (A.2) holds then $B \stackrel{d}{=} bA + a$ and a, b are the unique constants for which this holds.

Betekintés az extrémérték elméletbe

- X_1, X_2, \dots FAE változók
- $F(x)$ jelöli a közös eloszlásfüggvényt
- $M_n = \max(X_1, \dots, X_n)$ jelöli a maximumot
- $P(M_n < x) = P(X_1 < x, \dots, X_n < x) = F^n(x)$ a max elo. fv.-e
- $F(x)$ jobb végpontja (végtelen is lehet)

$$x_F = \sup\{x \in R : F(x) < 1\}$$

- Könnyen látszik, hogy 1 valószínűséggel $M_n \rightarrow x_F$

Milyen feltételek esetén vannak $a_n > 0$ és $b_n \in R$ sorozatok amellyekkel

$$\frac{M_n - b_n}{a_n} \xrightarrow{d} H,$$

ahol H nem degenerált

Theorem 3.2.3 (Fisher–Tippett theorem, limit laws for maxima)

Let (X_n) be a sequence of iid rvs. If there exist norming constants $c_n > 0$, $d_n \in \mathbb{R}$ and some non-degenerate df H such that

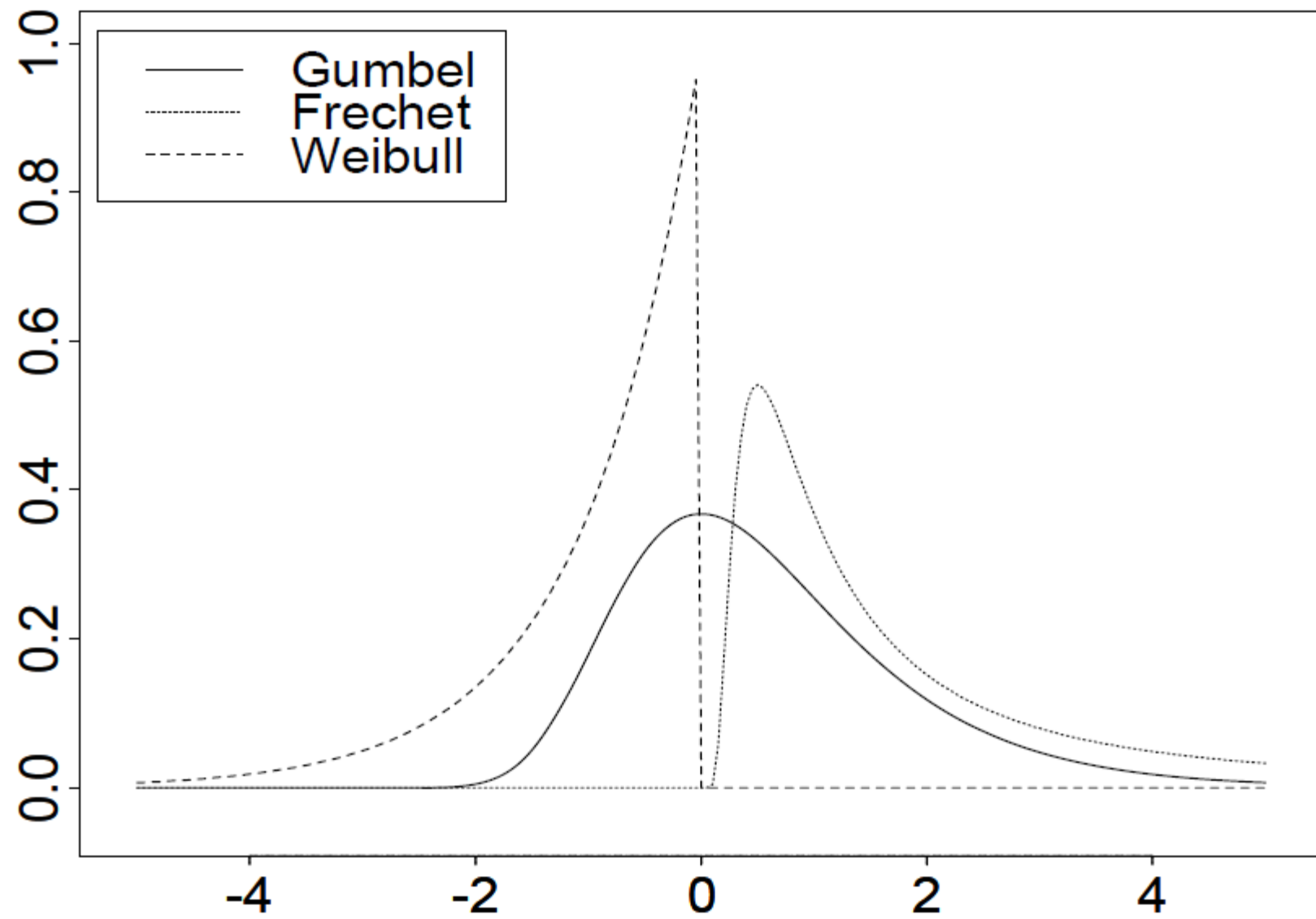
$$c_n^{-1} (M_n - d_n) \xrightarrow{d} H, \quad (3.9)$$

then H belongs to the type of one of the following three dfs:

$$\text{Fréchet:} \quad \Phi_\alpha(x) = \begin{cases} 0, & x \leq 0 \\ \exp\{-x^{-\alpha}\}, & x > 0 \end{cases} \quad \alpha > 0.$$

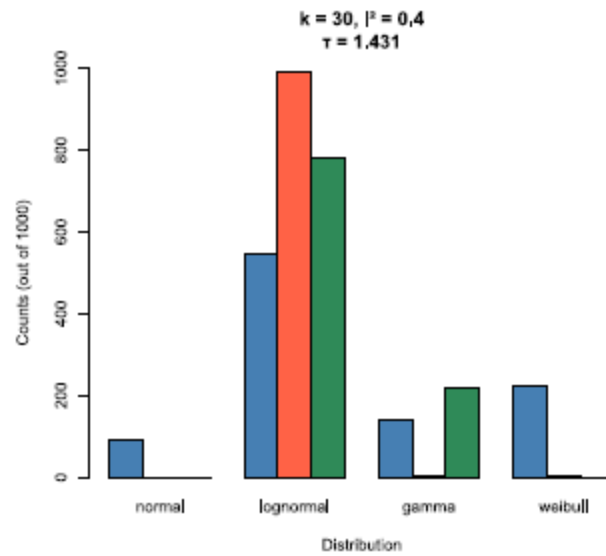
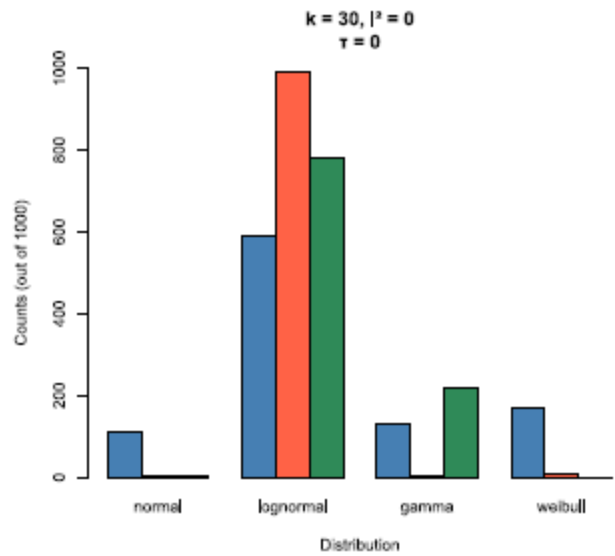
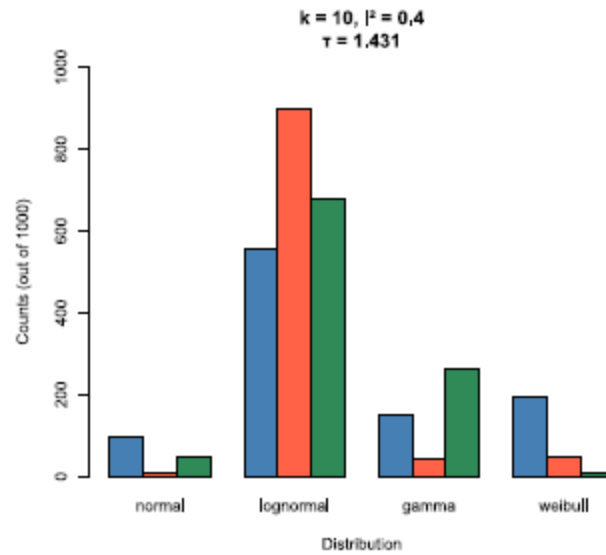
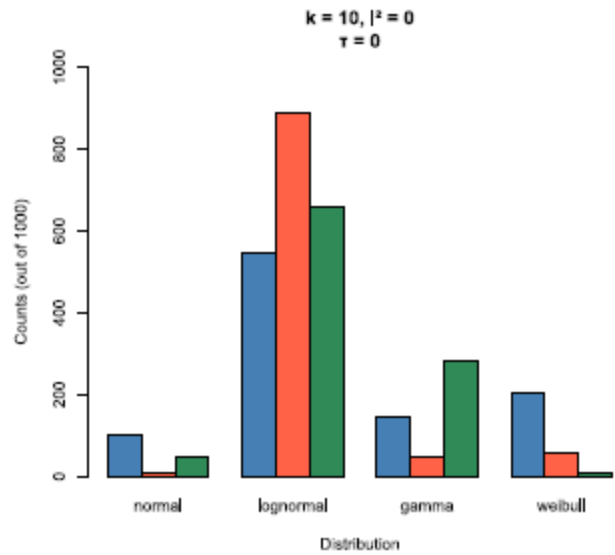
$$\text{Weibull:} \quad \Psi_\alpha(x) = \begin{cases} \exp\{-(-x)^\alpha\}, & x \leq 0 \\ 1, & x > 0 \end{cases} \quad \alpha > 0.$$

$$\text{Gumbel:} \quad \Lambda(x) = \exp\{-e^{-x}\}, \quad x \in \mathbb{R}.$$



Gumbel	$\Lambda(x) = \exp \{-e^{-x}\}, \quad x \in \mathbb{R}.$
MDA(Λ)	$x_F \leq \infty, \quad \bar{F}(x) = c(x) \exp \left\{ - \int_{x_0}^x \frac{g(t)}{a(t)} dt \right\}, \quad x_0 < x < x_F,$ where $c(x) \rightarrow c > 0, \quad g(x) \rightarrow 1, \quad a'(x) \rightarrow 0$ as $x \uparrow x_F$.
Norming constants	$d_n = F^{\leftarrow}(1 - n^{-1}), \quad c_n = a(d_n).$
Limit result	$c_n^{-1}(M_n - d_n) \xrightarrow{d} \Lambda$

Gamma	$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0, \quad \alpha, \beta > 0.$ $c_n = \beta^{-1}$ $d_n = \beta^{-1}(\ln n + (\alpha - 1) \ln \ln n - \ln \Gamma(\alpha))$
Normal	$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbb{R}.$ $c_n = (2 \ln n)^{-1/2}$ $d_n = \sqrt{2 \ln n} - \frac{\ln(4\pi) + \ln \ln n}{2(2 \ln n)^{1/2}}$
Lognormal	$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-(\ln x - \mu)^2/2\sigma^2}, \quad x > 0, \quad \mu \in \mathbb{R}, \sigma > 0.$ $c_n = \sigma(2 \ln n)^{-1/2} d_n$ $d_n = \exp \left\{ \mu + \sigma \left(\sqrt{2 \ln n} - \frac{\ln(4\pi) + \ln \ln n}{2(2 \ln n)^{1/2}} \right) \right\}$

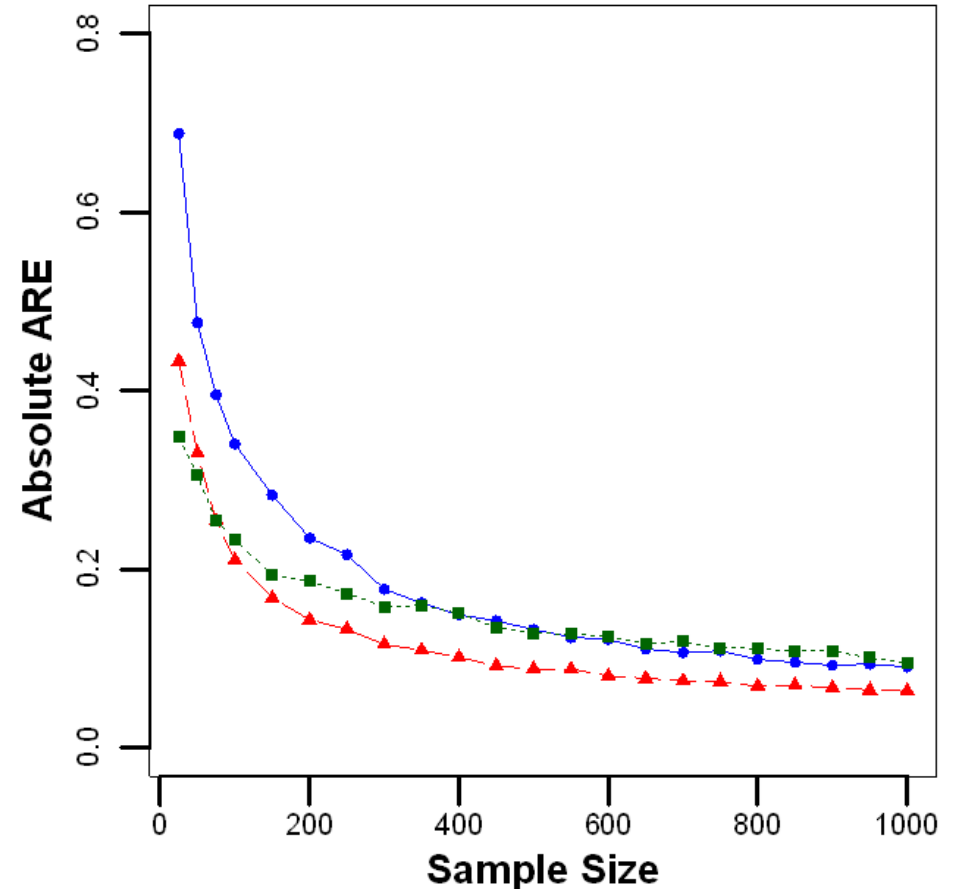


Folyamatban levő kutatás (nem része a megjelent cikknek)

Köszönöm a figyelmet!

Köszönet illeti

- Vető Bálintot az extrémérték elmélethez kötődő hasznos beszélgetések miatt
- NKFIH 142169 kutatási pályázatot a támogatásért
- Stipendium Hungaricum programot a PhD lehetőség biztosításáért
- A Semmelweis Egyetem Szaknyelvi Intézetét a nyelvi ellenőrzésért



ON THE DISTRIBUTION OF A VARIATE WHOSE LOGARITHM IS
NORMALLY DISTRIBUTED

By D. J. FINNEY

Rothamsted Experimental Station

1. *Introduction*

[Home](#) > [BMC Medical Research Methodology](#) > [Article](#)


Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan–Meier survival curves


Technical advance | [Open access](#) | Published: 01 February 2012

Volume 12, article number 9, (2012) [Cite this article](#)

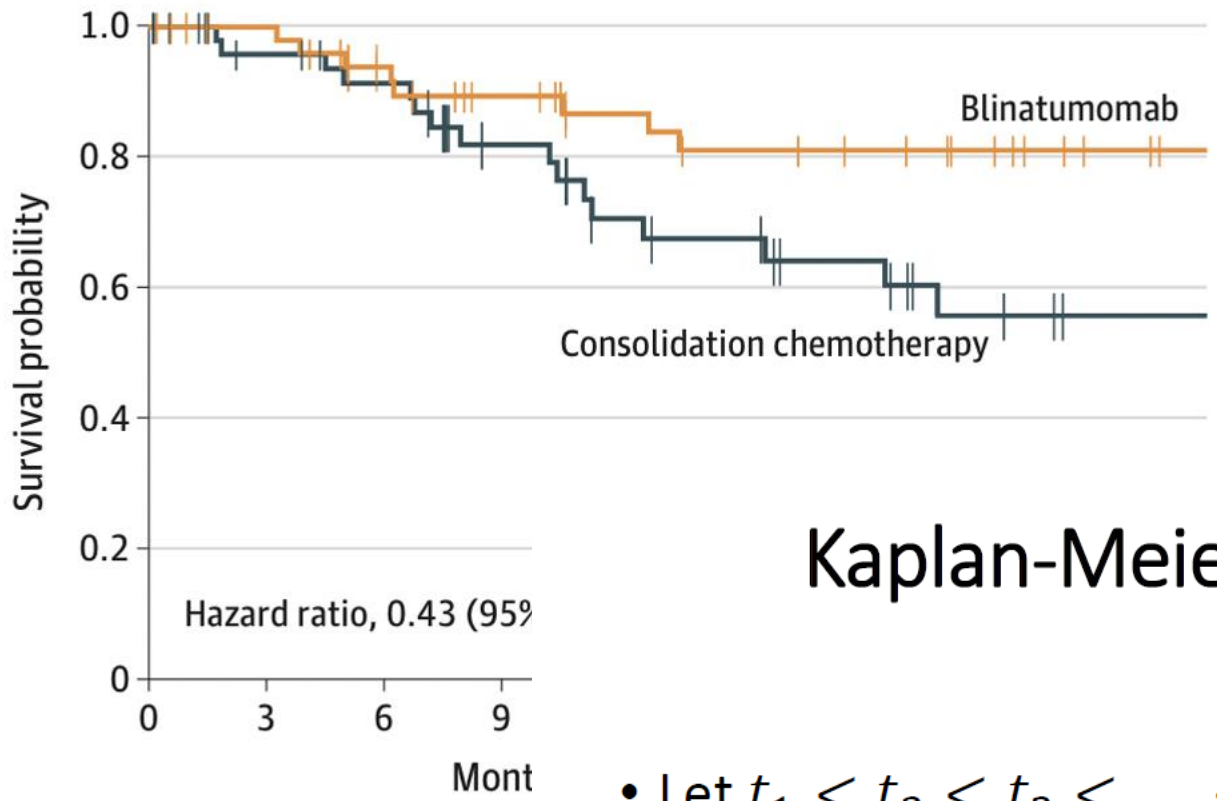
✔ You have full access to this [open access](#) article

[Download PDF](#) 

 [Save article](#)

[Patricia Guyot](#) , [AE Ades](#), [Mario JNM Ouwens](#) & [Nicky J Welton](#)

B Overall survival



No. at risk	0	3	6	9
Blinatumomab	54	50	42	36
Chemotherapy	54	45	41	30

Kaplan-Meier survival estimate for IntS-s

- Let $t_1 < t_2 < t_3 < \dots < t_k$ denotes the uncensored event times
- Let n_j denote the number of patients alive just before t_j
- Let d_j denote the number of events at t_j
- We calculate the survival probabilities at the uncensored event times using the following recursion:

$$\hat{S}(t_j) = \hat{S}(t_{j-1}) \left(\frac{n_j - d_j}{n_j} \right) = \hat{S}(t_{j-1}) \left(1 - \frac{d_j}{n_j} \right)$$

Visszabecsült IPD

TIME	EVENT
1.1	0
2.3	0
4.1	1
4.2	0
4.8	1
⋮	
⋮	

HA NINCS CENLORÁLIS
ALYKOR

• KLASZIKUS MINTA

• $KM \approx 1 - \underbrace{F_n^*(x)}$

EMP. ELOKLEJTV

APATIDOK (a) MEAN, SD

HELYETT (b) MEDIAN, Q₁, Q₃