

Statistika ajaloost maailmas ja Eestis

Ene-Margit Tiit

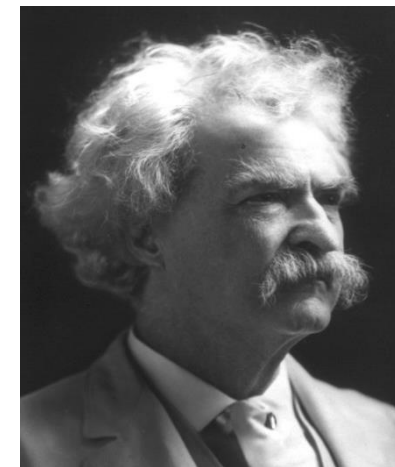
„Sissejuhatus erialasse“, 19.09.2019

Sisukord

- Statistika kui teaduse kujunemislugu
 - Statistika saab alguse rahvaloendustest
 - Rahvastiku loendamine ja rahvastikustatistika Eestis
 - Tartu Ülikoolist statistikateadusele XIX sajandil
 - Mõõtmise ja mõõtmisviga – statistikateaduse ehituskivid
 - Juhuslikkuse demüstifitseerimine. Tõenäosusteooria
 - Statistilised mudelid. Prognoosimine
- Statistika areng 20. sajandil
 - Matemaatiline statistika ja tõenäosusteooria Tartu Ülikoolis XX sajandil
 - Riikliku statistika areng Eestis XX sajandil
 - Matemaatiline statistika Matemaatika-informaatika- teaduskonnas.
 - Statistika ja andmeteadus

Statistika kui teaduse
kujunemislugu

Statistika vahekord tõe ja valega



Üks tuntumaid väljendeid statistika kohta on Mark Twaini tsitaat:

„On olemas kolme tüüpi valesid: valed, alatud valed ning statistika“.

Kui tõsiselt võetav see väide on?

- Ühelt poolt on tõepoolest võimalik mitte-asjakohase statistikaga valesid väiteid demagoogiliselt õigustada, kuid tavaliselt näevad asjatundjad selle nipi läbi.
- Teiselt poolt on aga hästi teada, et igasuguste väidete tõestamisel on empiirilistele andmetele tuginev statistika kõige võimsam vahend.

Selleks, et osata vahet teha õigete ja eksitavate väidete vahel, on tarvis tunda statistikat – eriti tänapäeval, kui nn libauudiseid levib palju.

Statistika saab alguse
rahvaloendustest

Rahvaloendused inimkonna ajaloos

- Vanimad statistika valdkonda kuuluvad tööd on rahvaloendused, mida on korraldatud juba enam kui 5000—6000 aastat tagasi kõigis vana tsivilisatsiooni maades – Lähis-Idas, Egiptuses, Hiinas, Kreekas.
- Rahvaloenduse tegemise eelduseks on piisavalt arenenud arvusüsteem, mis lubab loendada ka väga suuri arve ja nendega teha lihtsaimaid aritmeetilisi tehteid – liita ja lahutada.
- Pikka aega seisneski statistiliste andmete kogumine peamiselt loendamises. See oli kõikse statistika periood, mil ülesanded olid võrdlemisi lihtsad – peamiselt oli tarvis vastata küsimusele „Kui palju“ ja erimetoodikaid ei olnud tarvis luua.
- Rahvaloenduste teostamisel sai selgeks andmete tähtsus.

- Egiptuse vanimate rahvaloenduste toimumisajaks on hinnatud 3340–3050 eKr.
- Oletatavasti on Babüloni aladel leitud savitahvlitel kajastatud ka iidsete rahvaloenduste andmeid.
- Rahvaloendustest enne Kristuse sünni on teateid ka Hiinast, kus on dokumenteeritud ca 4000 aastat tagasi toimunud loendus.
- Äsja avaldas Kreeka statistik S. Missiakoulis artikli, milles kinnitab, et Ateenas toimus juba 16. sajandil enne Kristust rahvaloendus, mis oli tõenäoliselt vanimaid Euroopas: Ateena kuninga Kekropsi käsul viis iga linnakodanik teatavasse kohta linnaväljakul kivi. Nende kivide loendamise tulemusena tehti kindlaks, et Ateena elanike arv oli tollal 20 000.
- Rahvaloendusi korraldati ka Rooma riigis esimesel aastatuhandel enne Kristust, sealt pärineb rahvaloenduste rahvusvaheliselt käibiv nimetus (census).

2. P e a t ü t.

Kristusse sündimine ja ümberleikamine. Maria puhastamine. Simeon ja Hanna. Kristusse targad küsimised ja kostmised Jumala koidas.

Agga se sündis neil päivil, et Janna keisrist Augustussest väljaläks, et keif mailm pididi üleskirjotud sama.

2. (Sesinnane kirjotus sündis enne, kui Ririnius Siria ma üle waltitseja olli.)

3. Ja keif läksid endid lastma üleskirjotada, iggaüks omma linna.

4. Siis läks ka Josep Kalilea maalt Naatsaretti linnast üles Juda male Laweti linna, mis hütafs Betlemmaks, sepärrast et ta olli Laweti soust ja pärrismaalt,

5. Et ta ennast pididi lastma üleskirjotada Maria omma kihlatud naesega, kes olli käima peäl.

Matt. 1, 18.

Rahvaloendused Euroopas ja Eestis

- Euroopa riikides korraldati rahvaloendusi alates 11. sajandist (Inglismaal)
- Esimesed tänapäevased rahvaloendused toimusid 1719. aastal Preisimaal ja 1749. aastal Rootsis ja Soomes. Samal ajal toimusid loendused ka Kanadas ja USAs.
- Euroopa Statistikute I Kongressil 1865 lepidi kokku rahvaloenduste teostamise üldpõhimõtted.
- Eestis ja Lätis (Eestimaa, Liivimaa ja Kuramaa kubermangus) hakkasid loenduse korraldamist vedama Balti mõisnikud.
- 1860ndatel aastatel korraldati prooviloendusi mitmes Eesti piirkonnas
- 1881. a. loenduse küsimustiku koostas Tartu Ülikooli prof. Carl Schirren.

Rahvastiku loendamine ja rahvastikustatistika Eestis

CHARTE

VON

JENSEL.

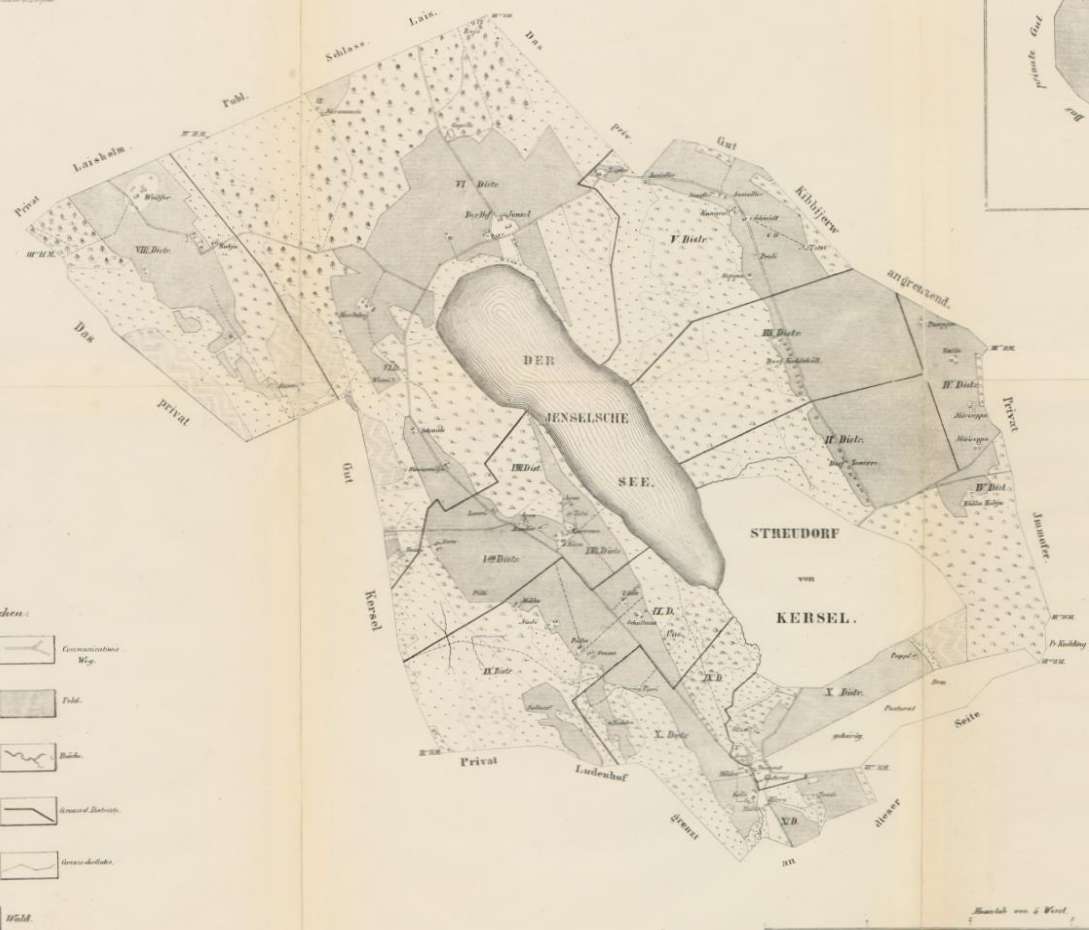
Angefertigt im Jahre 1864 von dem Ritterhofsbesitzer
Buschhond.

Leitung: v. C. Scholz in Dargatz.



Erklärung der Zeichen:

	Wasser		Commu. u. Privatw.
	Feld		Forst
	Wiese		Graben
	Moos		Bach
	Hügel		Zaun

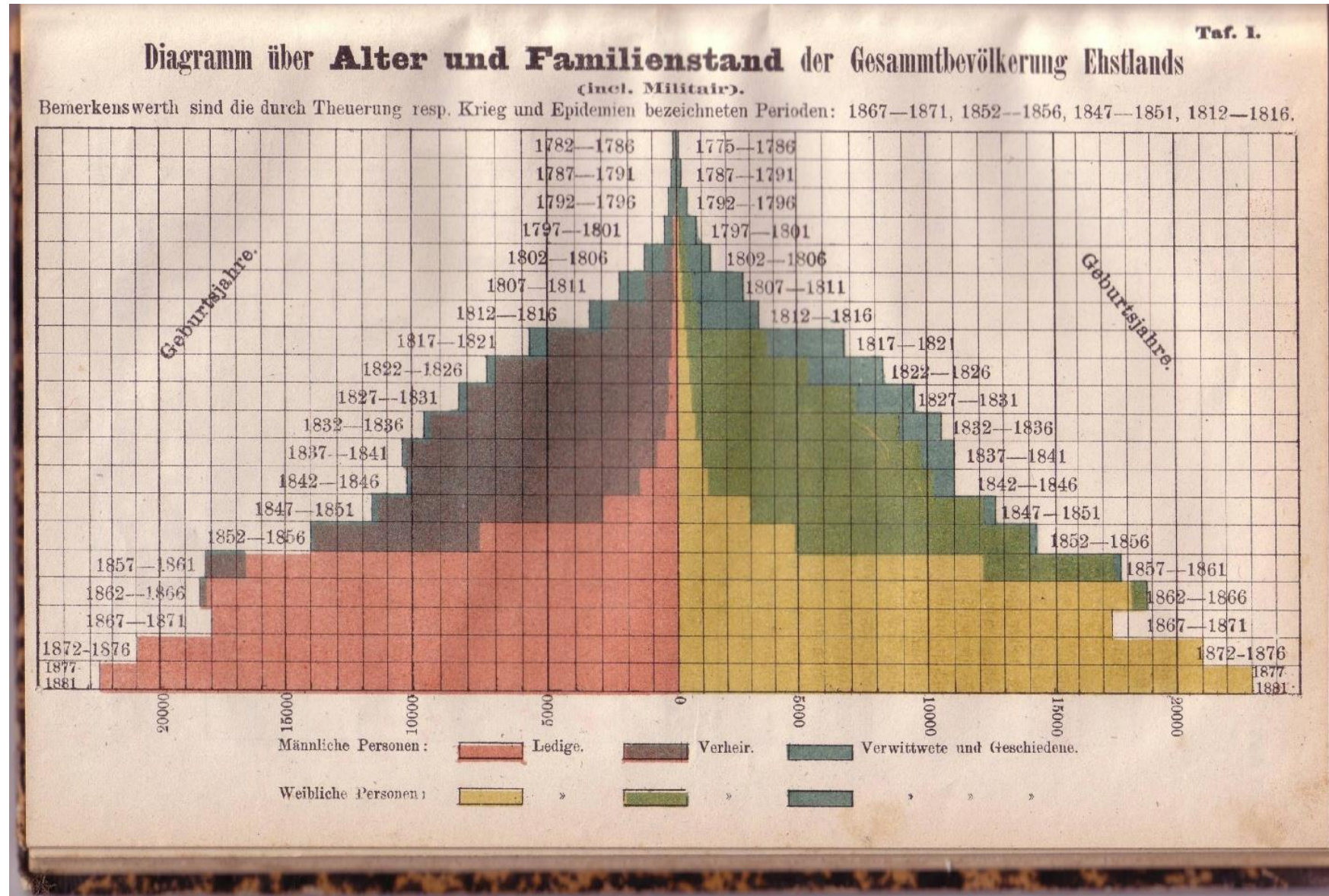


STREUDORF BONNIFER



Rahvaloenduse
kaart. Kuremaa,
1864.

Eesti rahvastikupüramiid 1881 (originaal)

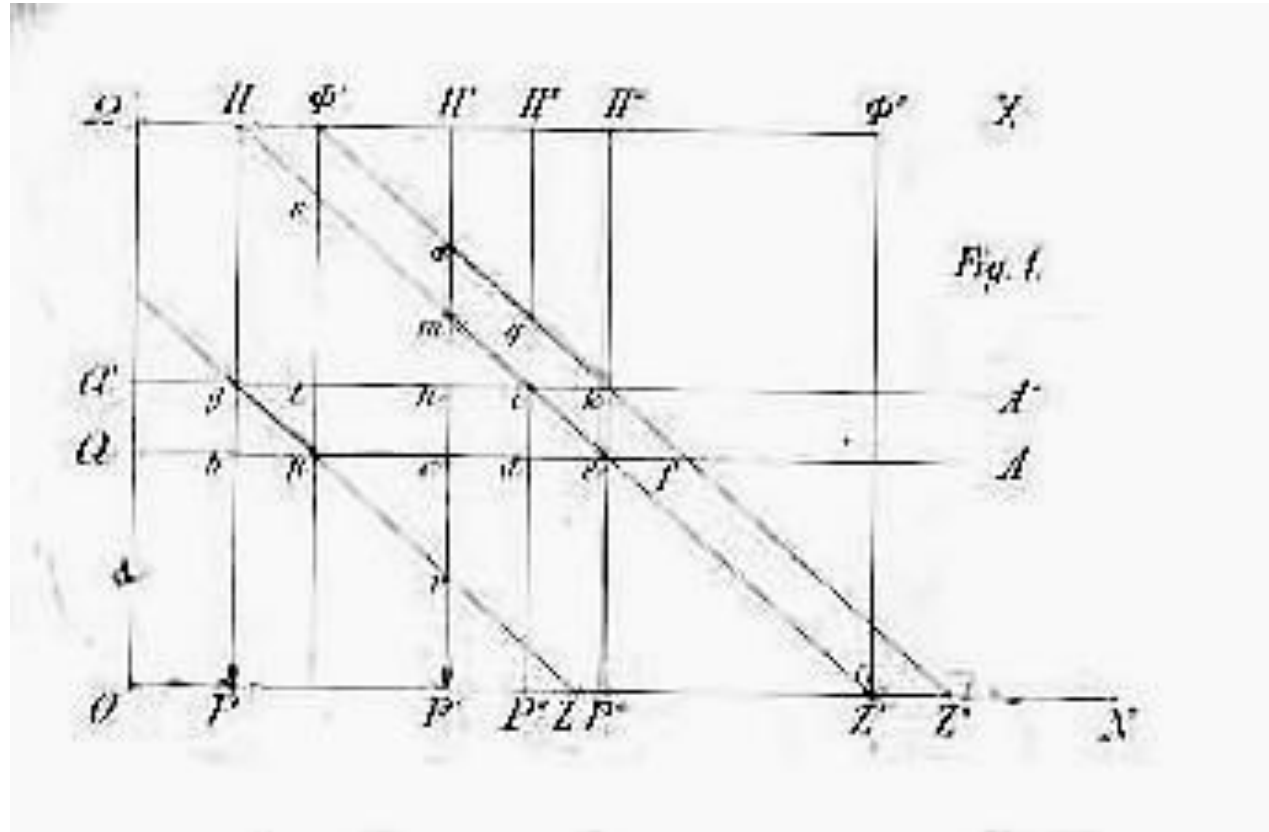
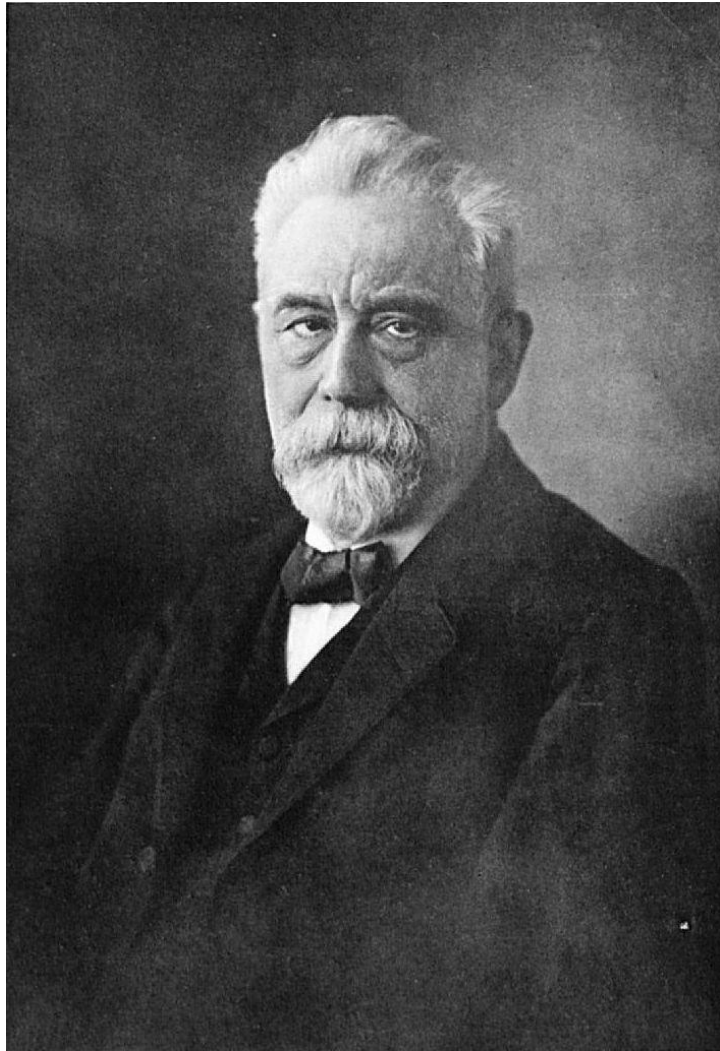


Tartu Ülikoolist statistikateadusele
XIX sajandil

Rahvastikustatistika alased uuringud Tartu Ülikoolis

- Kui Tartu Ülikool 1802. aastal taasavati, loodi siia kohe õppetool, mis esialgu kandis poliitaritmeetika nime, kuid juba 1804. aasta loengukursuste loetelus on kirjas nimetus **statistika**.
- Rahvastikustatistika alaste uurimustega tegutses **biostaatikute** töörühm (eesotsas Boris Körber).
- Kaitsti rida doktoritöid Lõuna-Eesti piirkonna rahvastiku kohta.

Wilhelm Lexis, Tartu Ülikooli Statistika professor
Introduction to the Theory of Population Statistics,
1875



Lexise diagramm, originaal

Mõõtmise ja mõõtmisviga –
statistikateaduse ehituskivid

- Astronoomia, samuti üks vanimaid teadusi, tekitas teist tüüpi andmeid.
- Loendamise kõrval oli tarvis ka mõõta – mõõdeti nurki/kaari taevalaotusel.
- Loendusandmetele lisandus uus andmete tüüp – mõõtmistulemused.
- Loendamist oli põhimõtteliselt võimalik teha täpselt, mõõtmistulemusele kaasneb juhuslik mõõtmisviga, seda püüti vältida ja vähendada.
- Sellega tegi seni üksnes loendamisele tuginenud olulise sammu teoreetilise matemaatika suunas.
- Juhuslikkusest sügavamalt arusaamiseks oli aga tarvis tõenäosuse mõistet.

Stonehenge, Inglismaa, 3500 a. e.Kr



Nebra, Saksamaa, 1600 a.e.Kr



Ajalooline astronoomia Eestis

- Eestis on leitud tehisaukudega kive, millel kujutatud suuremaid tähti ja tähekoogusid.
- Suurima panuse astronoomia ja geodeesia arengusse tegi Tartu Tähetorni astronoom Wilhelm Struve.
- Tema mõõdetud meridiaani kaar ja selle põhjal tehtud teaduslikud järeldused hämmastavad erakordse täpsusega



Juhuslikkuse demüstifitseerimine.
Tõenäosusteooria

Juhuslikkuse lihtsustatud mudel

Juhuslikkuse mõistmiseks ja demüstifitseerimiseks pakkusid vahendi hasartmängud, mida harrastasid ja uurisid mitmed nimekad matemaatikud alates 16. sajandist. Väga lihtsad, paljukordselt korratavad katsed (täringu, mündi ja kaardipaki abil) ja saadud tulemuste analüüs võimaldasid sõnastada seaduspärasusi.

- ❑ Juhuslikkus kaotas osa oma müstilisusest Blaise Pascali ja Pierre de Fermat kirjavahetuse tulemusena (1655).
- ❑ Juhuslikkuse mõistmist arendasid edasi Christiaan Huygens ja Jacob Bernoulli – esimeste tõenäosusteooria kui matemaatilise teaduse kohta kirjutatud raamatute autorid (17. saj II pool, 18. saj algus).



Blaise Pascal (1623—1662)



Pierre de Fermat (1601—1665)



Christiaan Huygens (1629-1695)



Jacob Bernoulli (1655—1705)



Statistilised mudelid.
Prognnoosimine

Kas tõenäosuslik mudel on rakendatav keerukamate andmete analüüsimisel?

- Tõenäosusteooria kui juhuslikkuse matemaatilise käsitlemise võtmeküsimus:
- Kas mõisted ja lahendused, mis sobivad hasartmängude jaoks, on rakendatavad ka reaalse elu ülesannete lahendamiseks?
- Vastuse andsid andmetele ja prognoosi mudelitele tuginevad teadusuuringud:
 - ❑ Inglase John Graunt'i 1662. aastal ilmunud uurimus "Loodusloolised ja poliitilised tähelepanekud (*observations*), mis põhinevad juurdelisatud materjalidel ning on tehtud surnute nimekirjade alusel" käsitles Londoni rahvastikku, sh suremust (elutabelid, epidemioloogia alged).
 - ❑ Graunti sõber William Petty avaldas samal aastal majandus-uuringutele pühendatud raamatu „Maksude ja panuste käsitlemine“.

Nende töödega pandi alus matemaatilisele statistikale, mis moodustab kõigi erinevate statistika suundade ja valdkondade ühtse teoreetilise aluse.

Vigade teooria ja vähimruutude meetod

- 18. saj lõpul ja 19.saj algul olid (eriti astronoomias) probleemiks ebatäpsed vaatlustulemused.
- Johann Carl Friedrich Gauss ja Adrien-Marie Legendre
 - töötasid sõltumatult välja mõõtmisvigade teooria,
 - ja lõid mudelite leidmiseks vähimruutude meetodi, et saada ebatäpsete mõõtmistulemuste põhjal võimalikult täpseid tulemusi.

Carl Friedrich Gauss (1777-1855)



Adrien-Marie Legendre (1752-1833)



Etienne Laspeyres, 1869–1873 [Tartu Ülikoolis](#), agregaatindeksite teooria looja
majandusstatistikas



$$\text{Laspeyres Index} = \frac{\sum (\text{Observation Price} \times \text{Base Qty})}{\sum (\text{Base Price} \times \text{Base Qty})}$$



Statistika areng 20. sajandil

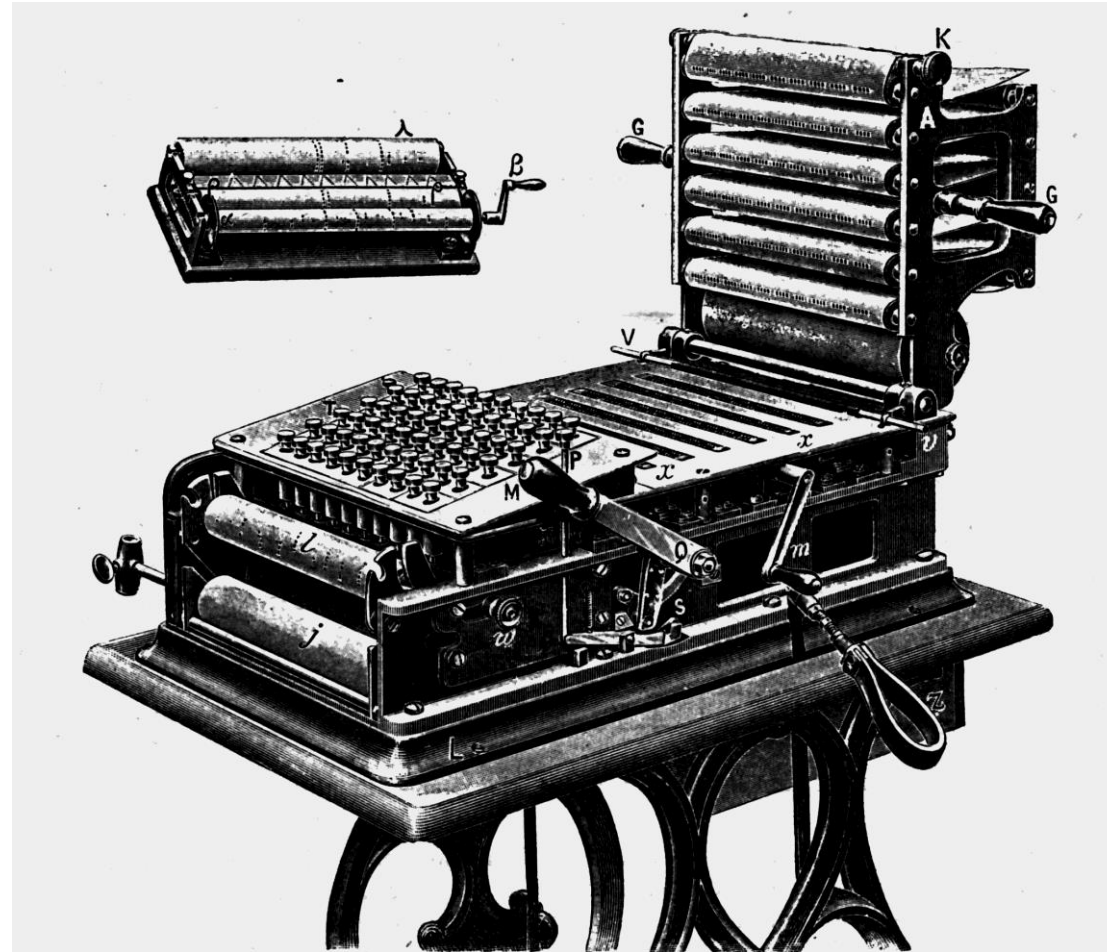
Hüpoteeside kontrollimine ja väidete tõestamine

- 20. sajandi künnisel ja esimesel veerandil intensiivistub empiiriliste andmete kasutamine paljudes teadustes.
- Toonaste vahendite kasutamisel on statistiliste mudelite rakendamine väga tõomahukas.
- Tekkis valikuuringute teooria, mis võimaldas osa kogutud andmete põhjal teha järeldusi terviku, so üldkogumi kohta.
- Statistika väljakutse oli – teha võimalikult õigeid järeldusi andmetest, mis sisaldasid juhuslikke vigu ja ühtlasi hinnata ka järelduse tõeväärtust.
- Sai selgeks, et statistika on möödapääsmatu vahend selleks, et tõestada teaduslikke hüpoteese katsete või empiiriliste mõõtmiste põhjal.
- Selle juures on oluline statistika erapooletu, sõltumatu roll mõõtmis-tulemuste käsitlemisel.

Statistika nõudis arvuteid

- Kuna statistika oli väga oluliselt seotud arvutustööga, siis on statistika arengut oluliselt mõjutanud arvutustehnika areng.
- Esimesed mehhaanilised arvutusmasinad loodi statistika tarbeks – nende abil töödeldi rahvaloenduse andmeid.
- See käib ka Eesti kohta – siia toodi Prantsusmaalt mehhaaniline arvutusmasin 1922. aasta rahvaloenduse andmete töötlemiseks. Ajavõit võrreldes käsitööga oli peaaegu kahekordne.
- Esimesed elektronarvutid, mis loodi II Maailmasõja päevil, lahendasid militaarotstarbelisi statistikaülesandeid.
- Arvutite sünnihetkest peale oli suur osa nende tööst andmete töötlemine – rahvaloenduste, ilmastiku, kvaliteedi kontrolli andmed.

Eestis 1922. aasta rahvaloendusel kasutatud arvutusmasin



Arvutuslükati. 1950ndate aastatel TÜ matemaatika tudengi tööriist



Matemaatiline statistika ja
tõenäosusteooria Tartu Ülikoolis
XX sajandil

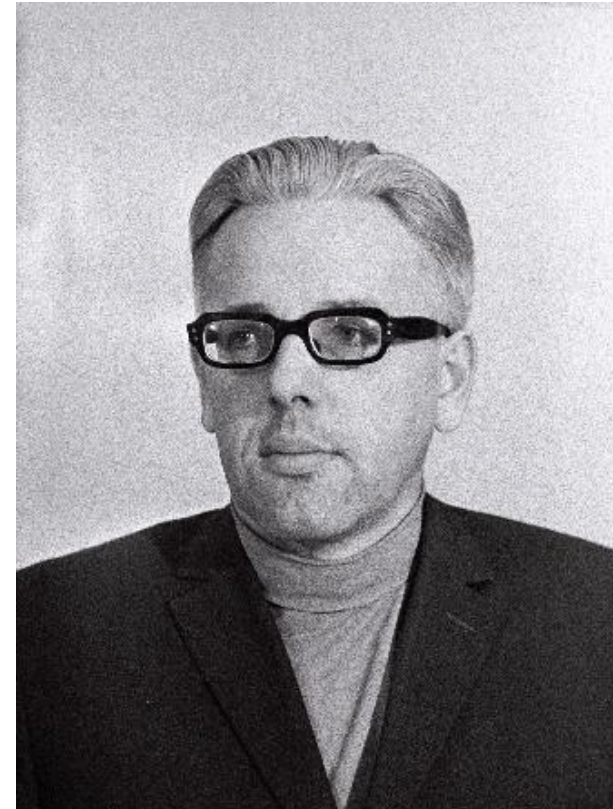
Matemaatiline statistika Eesti Vabariigi Tartu Ülikoolis

Prof. Gerhard Rägo, prof Arnold Humal,

dots. Edgar Krahm



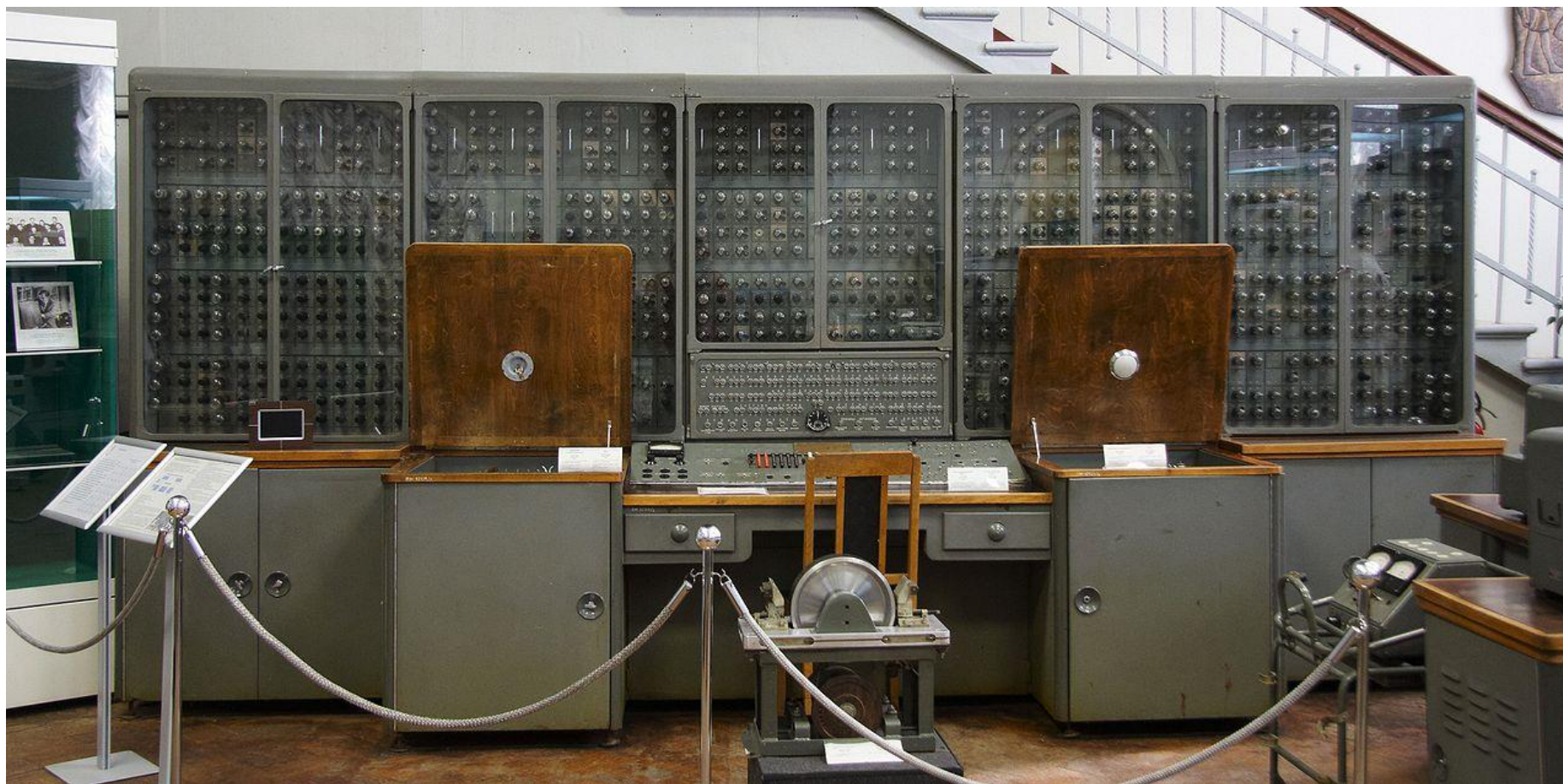
Statistikaõpetuse rajamine sõjajärgses Tartu
Riiklikus Ülikoolis ja Tallinnas TPIs
Leo Võhandu, Uno Mereste, Ivar Petersen



Rakendusstatistika ja andmeanalüüs Eestis

- Statistikat hakati Eestis intensiivselt rakendama paljudes teaduse valdkondades.
- Tartu Ülikoolis töötas aastail 1959—1989 arvutuskeskus, mis teenindas kõigi teaduskondade õppejõudude ja tudengite vajadusi statistiliste arvutuste järgi.

Ural-1 TÜ arvutuskeskuses, 1959



Statistika Eesti koolis

- Juba sõjaeelses Eesti Vabariigis jõudsid hariduse kujundajad arusaamisele, et koolides tuleb õpetada ka statistika aluseid. Selleks korraldati õpetajate täienduskursusi ja töötati välja õppematerjale, mis aga jäid okupatsiooni ja II Maailmasõja tõttu realiseerimata.
- Nõukogude korra tingimustes saavutasid Eesti koolide matemaatikaõpetajad õiguse ehitada koolimatemaatika kursus üles erinevalt üleliidulistest standarditest. Prof. Olaf Prinitza eestvõttel lülitati koolikursusesse statistika.
- Tänapäeval ei tähenda statistika mitte pikkade arvutuste tegemist, vaid sobiva tarkvara oskuslikku kasutamist. Selles suunas on tehtud paaril viimasel aastakümnel mitmesuguseid katsetusi.
- Üks läbimõeldumaid süsteeme on viimastel aastatel koos Wolframi instituudiga arendatav APS (Arvutipõhise statistika) programm, mille suundumus on statistika õppimine rakenduslike, erinevatest valdkondadest pärinevate praktiliste ülesannete najal.

Statistiline kirjaoskus

- Tänapäeval on arvuti koos elementaarsete statistikaprogrammidega igaühe laual.
- Kas see tähendab, et igaüks võib ja saab teha statistikat ja väitlustes statistilistele argumentidele toetuda?
- Statistika mõistmiseks on tarvis statistilist kirjaoskust!
- See tähendab järgmisi igapäevaelus vajalikke oskusi:
 - mõtestada statistilisi graafikuid ja neid ise joonistada/visandada,
 - mõista lihtsamaid statistikale tuginevaid väiteid, teha arvutustulemustest loogilisi järeldusi,
 - kasutada elementaarset statistika tarkvara.

On igati ootuspärane, et tänapäeval annab niihästi põhikool kui ka gümnaasium lõpetajatele elementaarse statistilise kirjaoskuse.



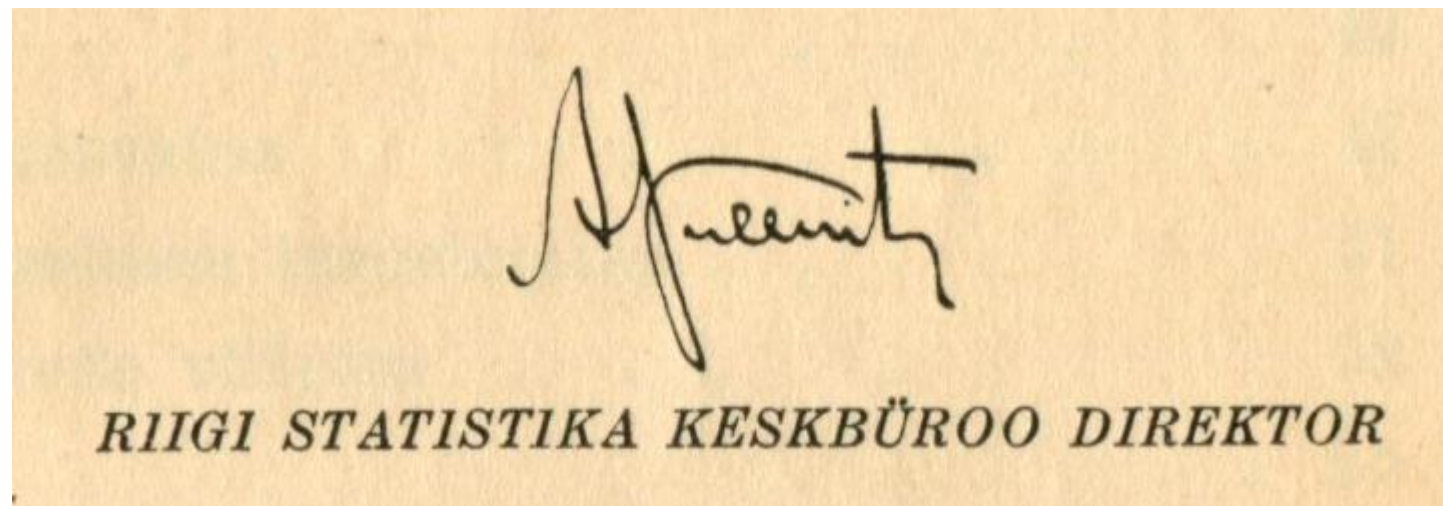
Riikliku statistika areng Eestis XX sajandil

Riiklik statistika Eesti Vabariigis

- Eesti Vabariigis asutati 1. märtsil 1921 Riigi Statistika Keskbüroo, mille juhiks sai [Albert Pullerits](#).



ALBERT PULLERITS



Eesti Vabariigi rahvaloendused

- Toimus kaks rahvaloendust: 1922 ja 1934.
- Koguti mitmesuguseid andmeid (rahvastiku, majanduse, kaubanduse, looduse jm kohta);
- Anti välja kuukirja ja aastaraamatuid;
- Tutvustati Eesti andmeid ka välismaale vastavate võõrkeelsete kogumike kaudu.

Riigistatistika okupeeritud Eestis

Nõukogude Liidu ja Saksamaa okupatsiooniga Eestis muutus oluliselt statistika tegemise olukord.

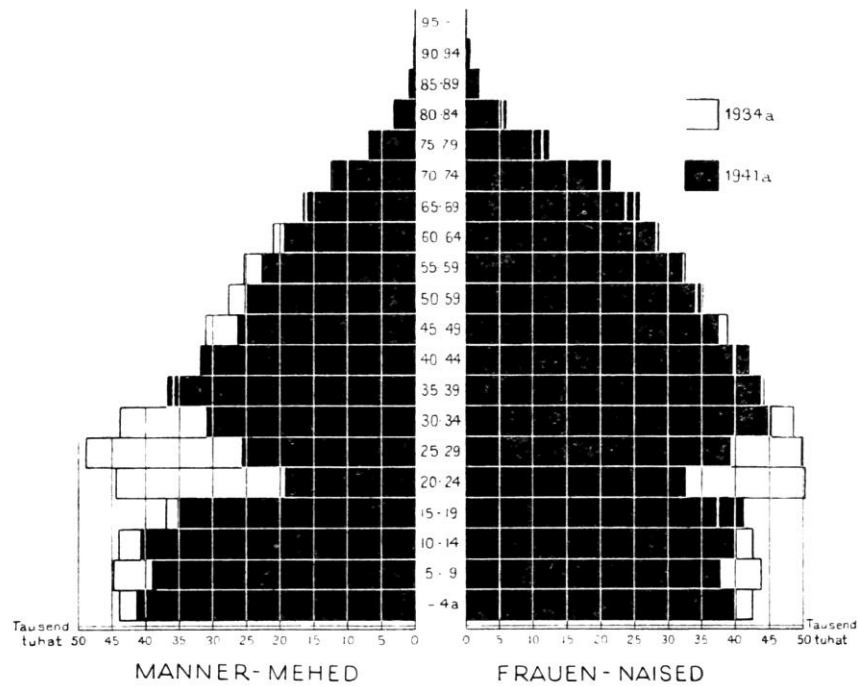
Statistika tegemine allutati otseselt okupeerivale võimule ja statistika oli suurel määral salastatud.

Siiski töötas statistikaametis, mis küll oli allutatud okupatsioonivõimudele, rahvuslikult mõtlevaid statistikuid, kes hoolitsesid Eestile oluliste andmete kogumise ja säilimise eest.

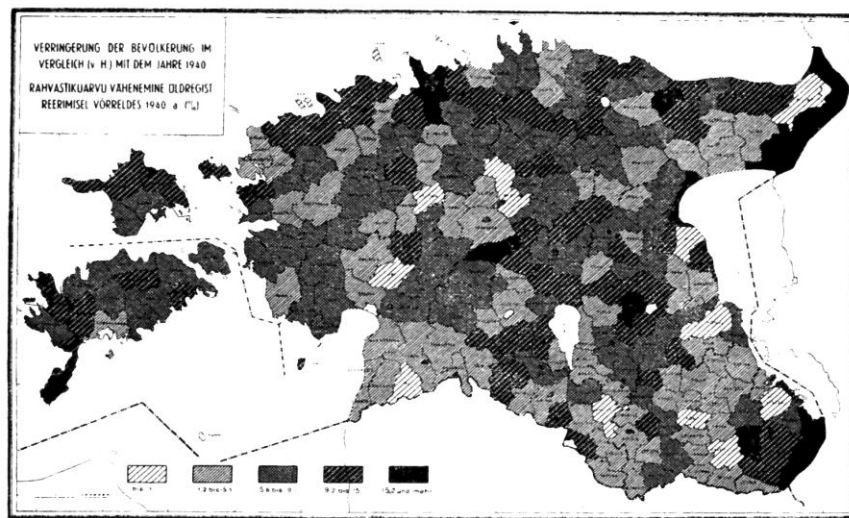
Kõige silmapaistvam oli selline tegevus rahvastikustatistika valdkonnas, mis oli rahvuse ja rahva säilimisloo seisukohast kõige olulisem statistikavaldkond.

Saksa okupatsioon

- Kolme aasta jooksul, sept.1941 – sept. 1944 korraldas Eesti statistikat Saksa kindralkomissarile alluv [Eesti Statistika Valitsus](#).
- Taas hakkas ilmuma Eesti Statistika Kuukiri, nüüd saksa- ja eestikeelsena ja märkusega „Ametkondlikuks kasutamiseks“.
- Praktiliselt kohe peale Eesti okupeerimist, 1.detsembril 1941 korraldati Eestis rahvaloendus. See oli väga napp (7 küsimust) ja tulemuste kohta on teada vaid 8 avaldatud tabelit.
- 1942. a. ilmus saksakeelne Eesti statistika aastaraamat
- Jätkus ka jooksev statistika kogumine, sh rahvastiku arvestus. Arvele võeti kõik sõjaväele potentsiaalselt vajalikud esemed – suusad, jalgrattad, parknahk jne.



Altersaufbau der Bevölkerung in 1934 und 1941
Eesti rahvastiku vanuspüramiid 1934. ja 1941. a. viieaastaste vanusrühmade järgi



Nõukogude okupatsioon

- Nõukogude võimu taaskehtestamise järel 1944. aasta lõpus hakkas Eesti statistikaelu korraldama [Eesti NSV Riiklik Statistikakomitee](#), mis allus Moskvas paiknevale Statistika Keskvalitsusele.
- Toonane statistikakomitee pidi olema valdavalt venekeelne nõukogude võimule kuulekas organisatsioon.
- Statistikakomitee põhieesmärk oli andmeid koguda ja Moskvasse saata, mitte neid kohapeal analüüsida ja tõlgendada.
- Statistiline aruandlus hõlmas suures ulatuses riiklike plaanide täitmise infot.
- Põhiline andmete kogumise metoodika oli toona loendus. Suur osa tehtavast statistikast oli kiirstatistika.

Rahvaloendused

- Nõukogude okupatsiooni ajal toimus Eestis neli rahvaloendust – aastatel 1959, 1970, 1979 ja 1989.
- Kogu Nõukogude Liidus oli loendusmetoodika ühesugune, loendatavatelt küsiti täpselt ühesuguseid küsimusi, mis loenduslehtedel olid vene keeles.
- Küsimuste tõlkimine loendatavatele oli loendaja mure, tema kandis ka vastused loenduslehele.
- Loenduse eel tehti ühiskondliku tööna kindlaks igal aadressil tegelikult elavad inimesed.
- Probleemiks olid loenduskaardid, mis olid sihilikult moonutatud, eriti militaarobjektide läheduses.

Taasiseseisvunud Eesti statistika-süsteemi ümberkujundamine

- Iseseisvat Eesti statistikasüsteemi hakati arendama vastavaks iseseisva riigi vajadustele juba nõukogude perioodi lõpuaastail.
- 1989. aastal loodi Eesti Vabariigi Statistikaamet.
- Aastal 1990 võttis [ENSV Ülemnõukogu](#) vastu [statistika-seaduse](#) ja sõlmiti esimene rahvusvaheline koostööleping Rootsi statistika keskbürooga.
- Eesti taasiseseisvumise järel algas Statistikaameti tegevuse vastavusse viimine rahvusvaheliste nõuetega.
- 2000. aastal toimud esimene rahva ja eluruumide loendus taasiseseisvunud Eestis.
- Loendustulemused on lisaks paberraamatutele avaldatud ka veebis



STATISTIKAAMET

ISIKULEHT

3784751



V

Kinnitatud Vabariigi Valitsuse 5. märtsi 1999.a. määrusega nr 82
ANDMEID KASUTATAKSE AINULT STATISTILISTEL EESMÄRKIDEL

V

Ringkond	Piirkond	Jaoskond	Loendamiskoha eluruumilehe nr	Rea nr	<input type="checkbox"/> Järelküsitus
----------	----------	----------	-------------------------------	--------	---------------------------------------

1. Eesnimi	
2. Perekonnanimi	
3. Sugu	<input type="checkbox"/> mees <input type="checkbox"/> naine <input type="checkbox"/> Vastab isiklikult
4. Sünniaeg (PP KK AAAA)	5. Isikukood

6. Püsielukoht		
a) <input type="checkbox"/> Loendamiskoha eluruumis		
Püsielukoha eluruumilehe nr	Rea nr	
b) Eluruumis, mis asub väljaspool loendamiskoha jaoskonda ja kus loenduse ajal kedagi ei viibi		
c) Täidetakse väljaspool loendamiskoha jaoskonda paikneva eluruumi puhul, kus loenduse ajal keegi viibib:		
tänav (talv)	maja nr	korteri nr
asula		
vald	maakond	
d) Välisriigis elaval isikul märkida riik:		
<input type="checkbox"/> Venemaa <input type="checkbox"/> Soome <input type="checkbox"/> Rootsi	muu riik (kirjutada)	

7. Kus Te viibisite ööl vastu 31. märtsi käesoleval aastal?		
a) <input type="checkbox"/> Püsielukohas		
b) <input type="checkbox"/> Loendamiskoha eluruumis, mis ei ole püsielukoht		
c) Täidetakse isikul, kes viibis kohas, mis ei ole tema püsielukoht ega loendamiskoht:		
tänav (talv)	maja nr	korteri nr
asula		
vald	maakond	
d) Välisriigis asunud isikul märkida riik:		
<input type="checkbox"/> Venemaa <input type="checkbox"/> Soome <input type="checkbox"/> Rootsi	muu riik (kirjutada)	

8. Milline on Teie kodakondsus?			
<input type="checkbox"/> Eesti	<input type="checkbox"/> Ukraina	<input type="checkbox"/> Läti	<input type="checkbox"/> Soome
<input type="checkbox"/> Vene	<input type="checkbox"/> Valgevene	<input type="checkbox"/> Leedu	<input type="checkbox"/> määratlemata
1. muu kodakondsus (kirjutada)	2. muu kodakondsus (kirjutada)		

V

3

V

Eesti NSV riikliku statistika käilakuju Lembit Tepp (1930—2008)

- TPI majandusteadlase haridusega Lembit Tepp töötas Eesti statistikasüsteemis alates 1961. aastast, olles mitme nõukogudeaegse rahvaloenduse peamine korraldaja või sisuline juht.
- Tähelepanuväärne on tema hoolikus rahvastikuandmete kogumisel ja säilitamisel ka siis, kui neid ametlikult ei avaldatud.
- Lembit Tepp on kirjutanud rea analüütilisi artikleid rahvastikustatistika kohta ja võtnud sõna ka ajakirjanduslikes väitlustes. .



17.10.2019

Esitluse või esitleja nimi



Matemaatiline statistika Matemaatika- informaatika- teaduskonnas.

Otsingud ja enesetõestus

*Ene-Margit Tiit
Märt Möls*

**Matemaatikateaduskonna 40. juubelit tähistav
konverents, 14. septembril 2007**



Väikese kateedri/instituudina on MSI oma tegevusaja jooksul korduvalt pidanud ennast ja enda vajalikkust tõestama. See on põhjustanud otsinguid ja tegevuse mõtestamist. MSI tegevuse aluseks on olnud.

- Orientatsioon ajakohasele ja ammendavale eestikeelsele statistikaõppele koos sellest tuleneva vajadusega arendada terminoloogiat ja luua õppekirjandust.
- Teadustöös tasakaal teooria ja praktika vahel, eesmärgiga teha uuringuid ühiskonnale olulistes suundades ja maailmas tunnustatud tasemel.
- Avatus ja aktiivsus suhtluses koostööpartneritega nii ülikoolis, Eestis kui ka välismaal
- Tugev meeskond, mis seob erinevate põlvkondade ja erinevate suundade inimesi.



1. Sünnilugu. Matemaatilise statistika kateedri ja instituudi kujunemine



Tartu Ülikoolis on selle taasloomisest saadik olnud üks või isegi mitu statistika professuuri ning siin on töötanud nimekaid statistika teoreetikuid (W. Lexis, E. Laspeyres), siin kaitsti maailma esimene matemaatilise epidemioloogia alane doktoritöö (P.D. En'ko).

Ometi ei olnud siin Matemaatika teaduskonna loomise hetkel 40 aastat tagasi ühtegi allüksust, mille nimest võinuks järeldada, et siin õpetatakse ka tõenäosusteooriat ja matemaatilist statistikat. Suuremas osas Euroopa ülikoolidest sellised üksused (instituudid, õppetoolid) sel ajal olid tekkinud või tekkisid kas matemaatika või mõne teise teaduskonna (majandus-, bioloogia, filosoofia-) teaduskonna juurde.

Matemaatilise statistika olulisuse mõistmise eest oleme tänulikud professoritele Ülo Kaasikule ja Leo Võhandule, pidades selles seoses meeles ka äsja manalasse lahkunud Ivar Peterseni.



Matemaatilise statistika ja programmeerimise kateeder 1970

Mihhail Lanin, Rein Tammeste (1939—1973), Ivar Kull (1947—1953), Tõnu Möls, Liina-Mai Tooding, Mare Koit, Ülo Kaasik (1926—2017), Ene-Margit Tiit, Tiina Lasn



Professor Taivo Arak

- Aastail 1985—1990 juhatas Matemaatilise statistika kateedrit Leningradi Ülikoolis doktoritiitli omandanud ja Markovi väljade uurimise eest maineka auhinna võitnud professor Taivo Arak (1946—2007). Tartust lahkus Arak Tallinna Tehnikaülikooli ja sealt peatselt Göteborgi Ülikooli .



Matemaatilise statistika instituut 1993

Ülo Vilismäe, Tarmo Koll, Kalev Pärna, Ene Käärik, Krista Fischer, Tõnu Möls, Kuldev Ääremaa, Imbi Traat, Anne-Mai Parring, Elvi Ehasalu, Ene-Margit Tiit, Sade Koskel (1938—2013), Mare Vähi



- Statistikutel on kogu tegevusperioodi jooksul õppetöös osalenud üle kahe korra rohkem inimesi kui on koosseisulisi õppejõude-teadureid.
- Õppetöösse on haaratud kraadiõppurid, kuid statistikute meeskonda on alati kuulunud ka selliseid inimesi, kes põhilise leiva teenivad mujal.



MSI ajaloos aitavad meid orienteeruda alates 1992. aastast igal aastal ilmunud brošüürid **Activities**, mille eesmärgiks on tutvustada meie tegemisi kolleegidele välismaal.



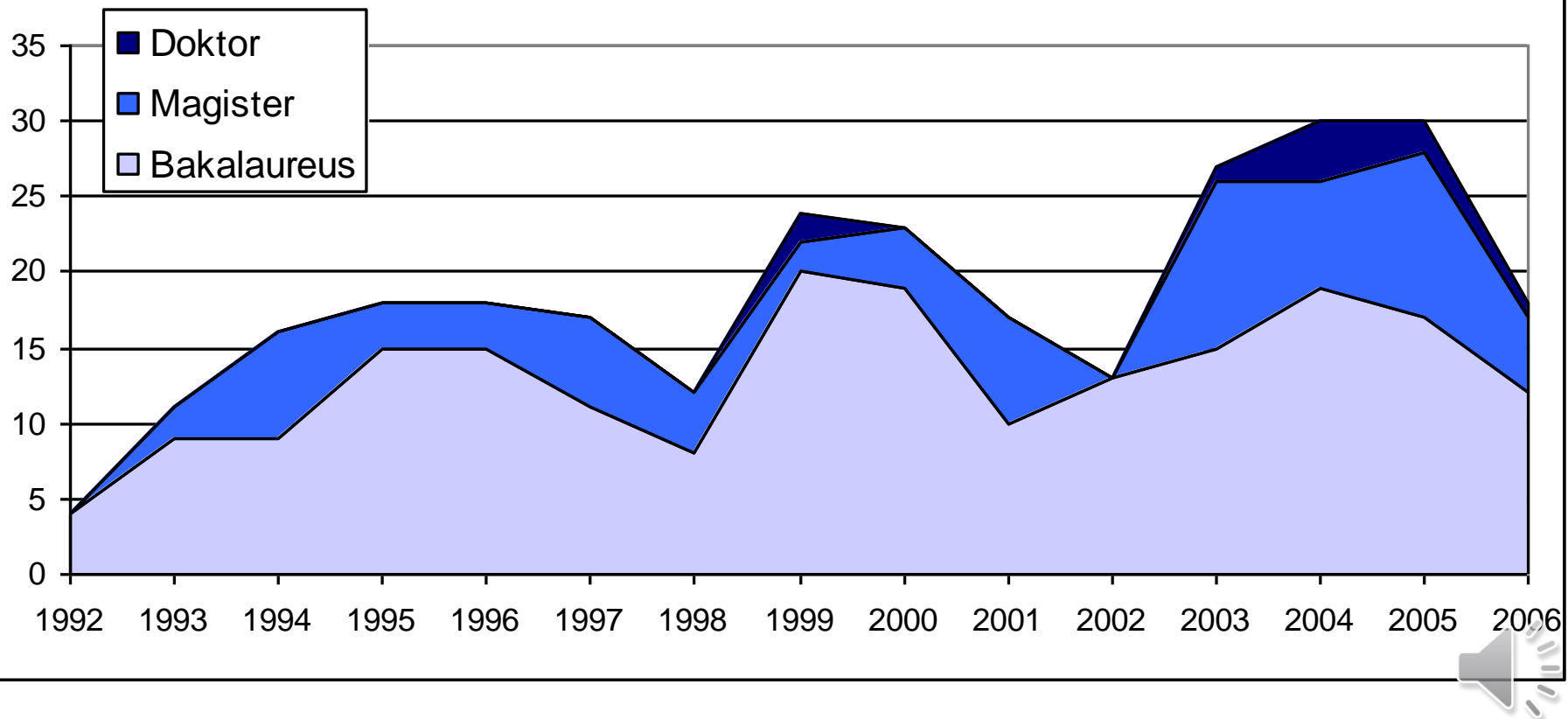
- Oluline osa matemaatilise statistika õppejõu tööst toimub esiisade viisil kriidiga tahvlil, nagu näitab prof Kalev Pärna



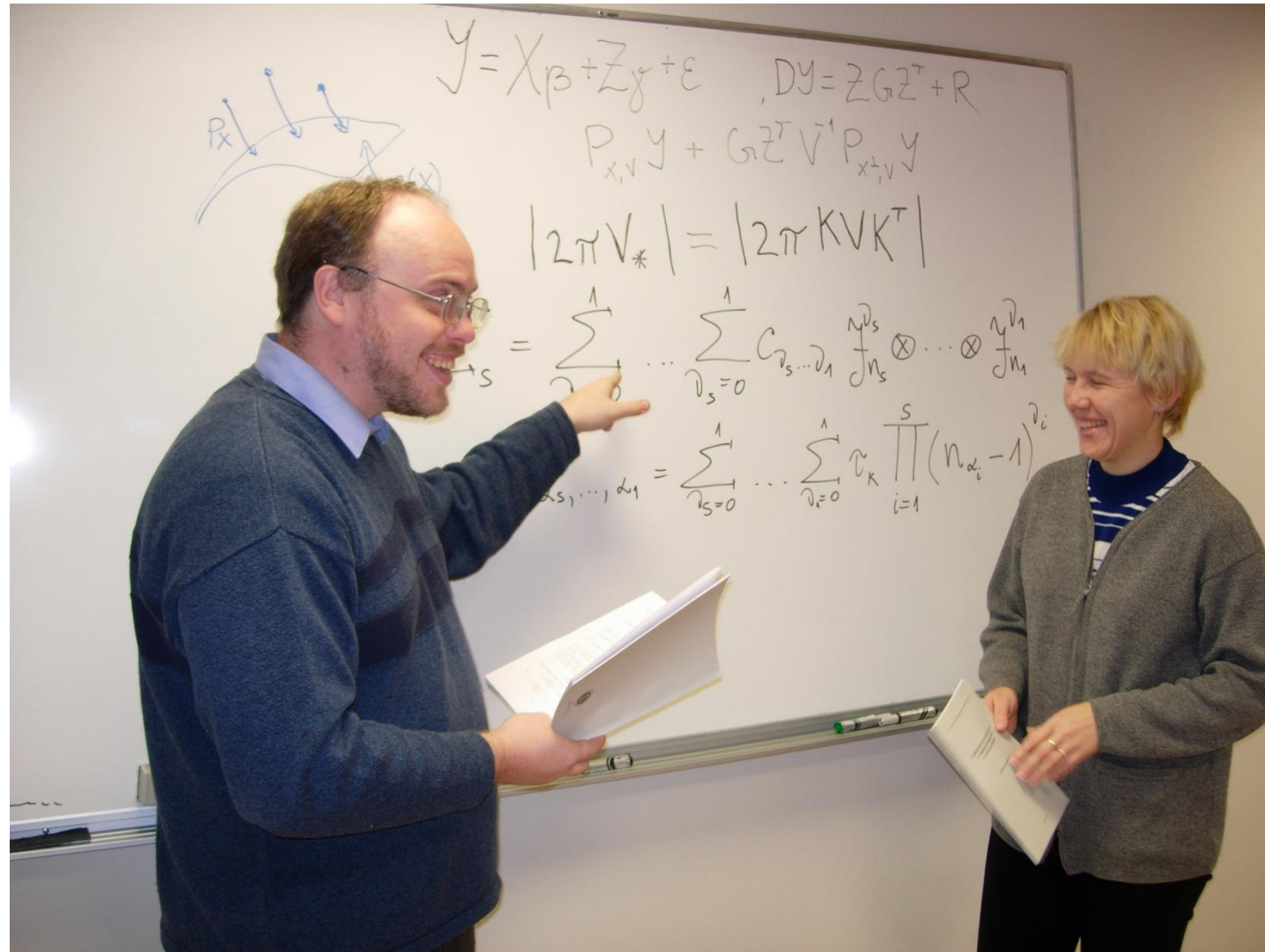
- Eesmärgist tulenevalt on õpetatavate kursuste arv kiiresti kasvanud, kuigi on jäänud ära kursuste keeleline dubleerimine (nt venekeelsed loengud füüsikutele ja ingliskeelsed meedikutele).
- Järjest väheneb teistes teaduskondades õpetatavate ainete osakaal. Üks põhjusi on siin meie õppejõudude üleminek – L.-M. Tooding töötab Sotsiaal- ja K. Fischer Arstiteaduskonnas.



Erineva tasemega lõpetajad aastate lõikes



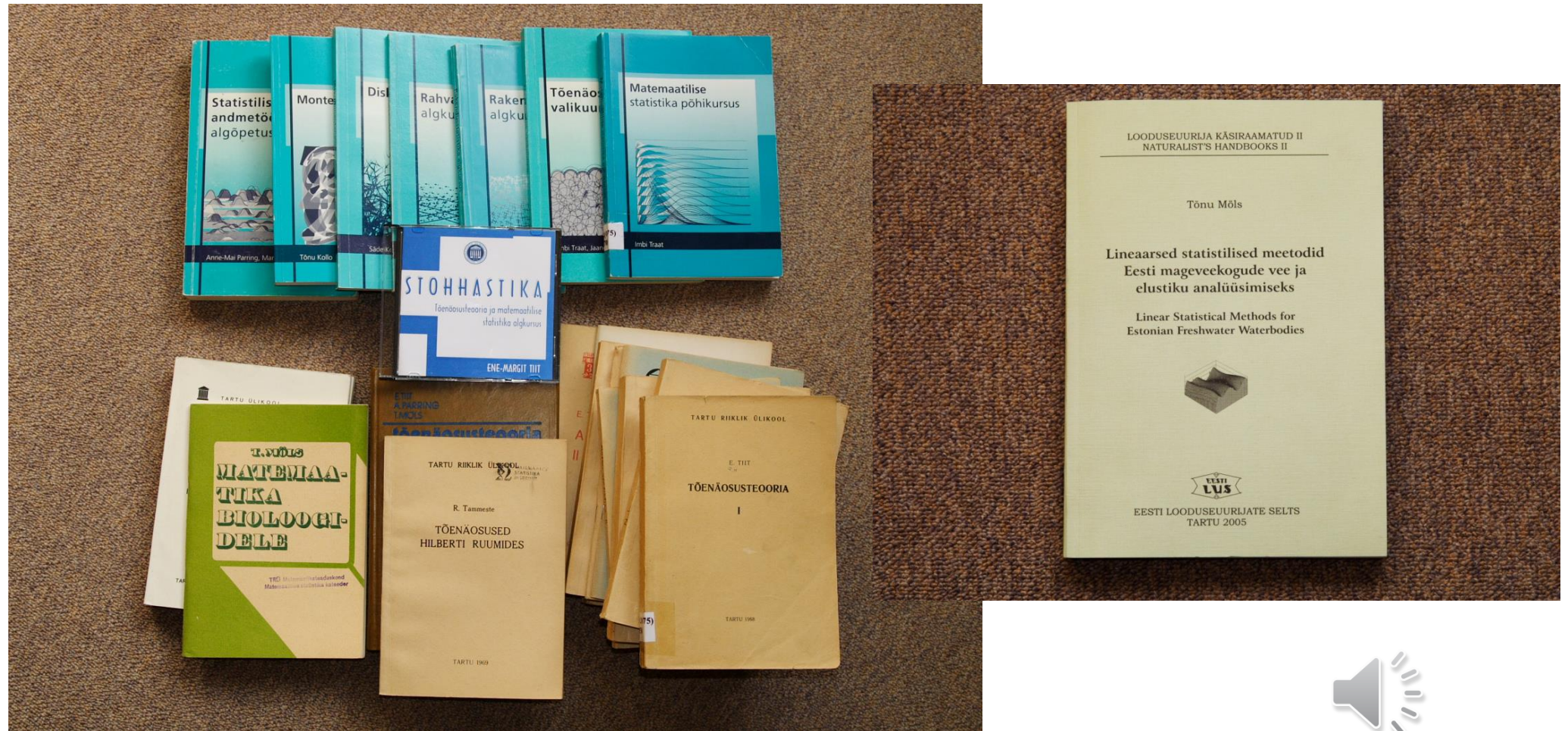
Seminaris lahendatakse probleeme üheskoos. Äsja doktoritöö kaitsnud Märt Möls ja Tatjana Nahtman (von Rosen)



- Keskmiselt lõpetab matemaatilise statistika erialal põhiõppe 75,4% sisseastunudest,
- Magistritöö kaitsjaid on 36,7% põhiõppe lõpetanutest;
- Doktoritöö kaitsjaid on 18,1% magistritöö kaitsjatest



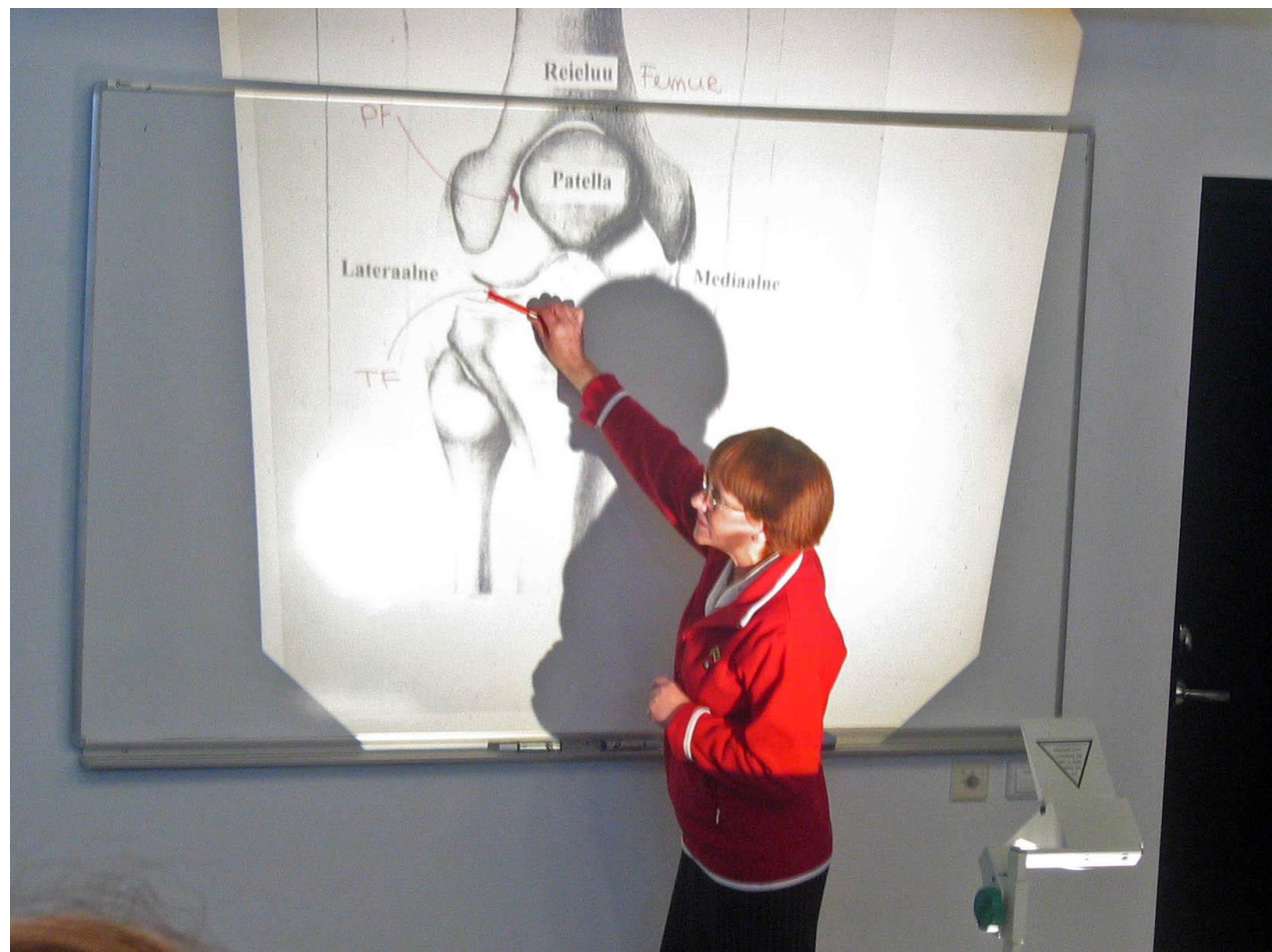
MSI õppejõududel tuli alustada tööd eestikeelsete õpikute kirjutamise ja ka oskussõnavara väljatöötamisega



- Neljakümne aasta jooksul MSI õppejõudude poolt kirjutatud ja avaldatud ligi 40 originaalset õpikut, kokku ligi 8000 lk.
- Viimasel ajal on siiski paberkandjal õpikute kirjutamise aktiivsus vähenenud, neid asendavad üliõpilastele küttesaadavad õppematerjalid internetis.
- Muutunud on ka loengud: väga sageli on abivahendiks arvutiprojektor esitamaks statistilisi andmeid, skeeme, jooniseid ja arvutustulemusi.



Statistikul tuleb sageli lahendada reaalse eluga seotud ülesandeid: dots. Imbi Traat kontrollib meditsiinist tulenevaid statistilisi hüpoteese ja demonstreerib saadud tulemusi.



Statistika tarkvara

Statistika õpetamise juures on MSI õppejõud käinud läbi arvutustehnika arengu perfokaartidega toidetavast suurarvutist tänapäevase süle-arvutini.

Oleme hoolitsenud ka statistika tarkvaraga seotud eestikeelse õppekirjanduse eest.



Statistika oskussõnavara

- Uute ainete õpetamisel tekib vajadus luua uusi omakeelseid oskussõnu. Paljud neist, nagu valim, hinnang, nihe jne on saanud üldtuntuiks.
- Eesti statistikud on oma panuse andnud ka **ISI (Rahvusvahelise Statistikainstituudi) paljukeelse sõnastiku** loomisele. Selle vahendusel on võimalik leida eestikeelseid vasteid 28s keeles esitatud statistika-terminitele (<http://isi.cbs.nl/glossary.htm>).





Teadustöö Matemaatilise statistika instituudis



Mitmemõõtmeline statistiline analüüs

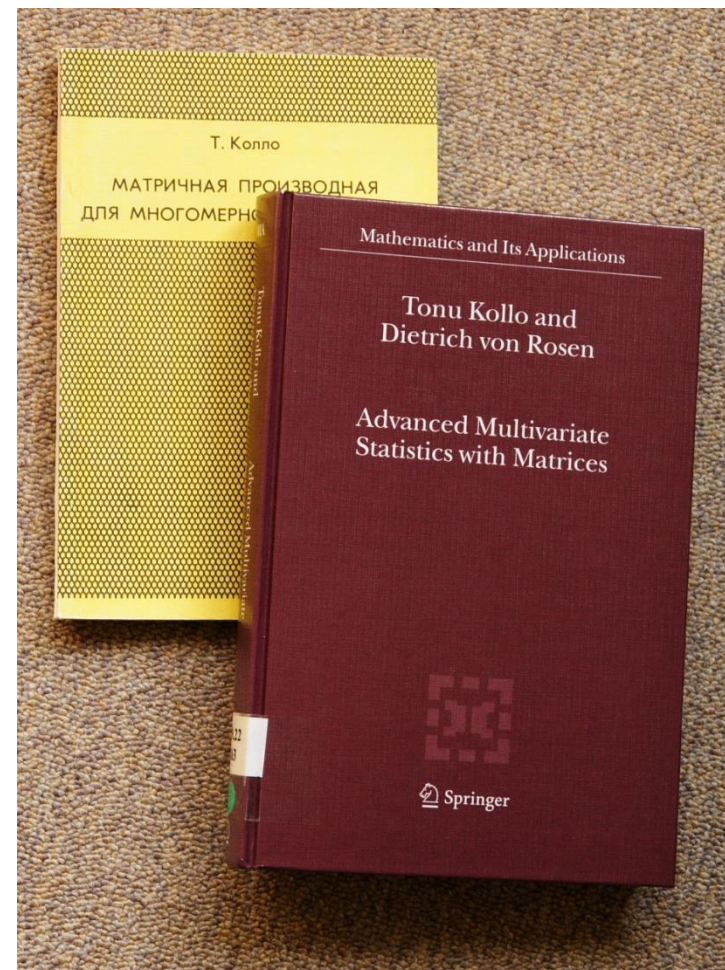
Tartus tehtav matemaatilise statistika alane teoreetiline uurimistöö kuulub olulises osas mitmemõõtmelise statistika valdkonda.

Selleks, et adekvaatselt kirjeldada reaalses maailmas, on tarvis arendada ja uurida teoreetilisi mudeleid, mis võimaldavad analüüsida paljude juhuslike suuruste koostoimimist ja vastastikuseid mõjutusi.

Mitmemõõtmelise statistika mudelite arendamiseks on tarvis edasi arendada maatriksaparatuuri.



Tõnu Kollo on riikliku teaduspreemia laureaat
aastast 2006



Matemaatilise statistika instituudis on 1998—2007 a kaitstud 10 doktoritööd, lisandub 3 välismaal kaitstud tööd



Tõnu Möls ja üks tema edukatest õpilastest – doktor Tanel Kaart



Doktoritööd kaitseb Meelis Käärik (juhendaja
prof K. Pärna), jälgivad dr Jüri Lember ja
doktorant Ants Kaasik



Statistikute publikatsioonid

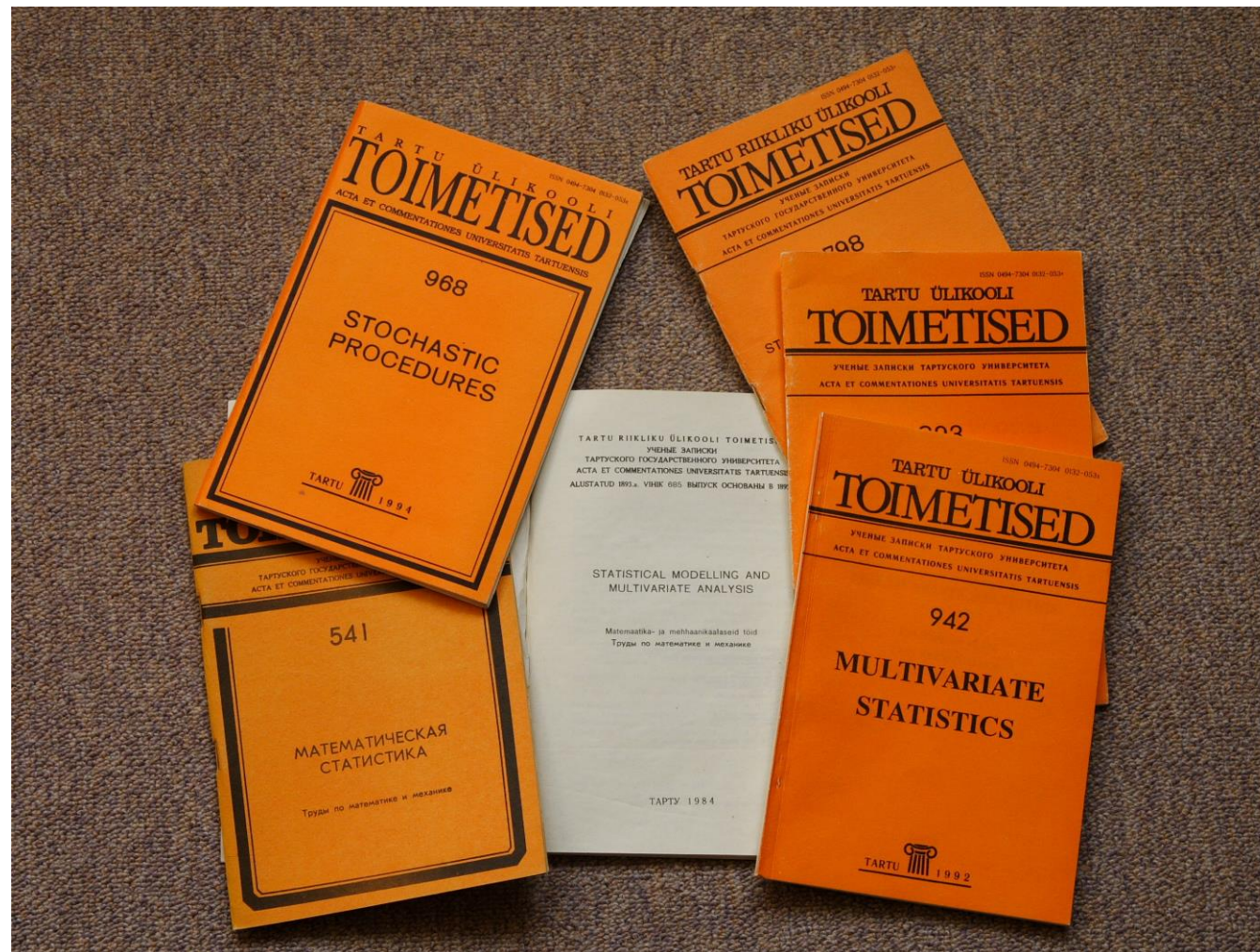
- Alates 1968. aastast kuni 1992 aastani ilmus statistikute sulest 15 venekeelset kogumikku “Arvutuskeskuse Töid” (kokku 10,5 cm) ja 8 eestikeelset “Programme kõigile” (5,2 cm)
- MSI esimene ingliskeelne kogumik (***Acta Universitatis Tartuensis***) ilmus 1984. aastal. Seni on ilmunud 6 kogumikku.
- Käesoleval sajandil avaldatakse teadustulemusi valdavalt rahvusvahelistes ajakirjades/ kogumikes



Arvutuskeskuse väljaanded



Acta Universitatis Tartuensis statistikaalased vihikud





4. Statistikute avatus maailmale

- Osalus rahvusvahelistel konverentsidel
- Rahvusvaheliste konverentside korraldamine Tartus
- Tudengite ja õppejõudude akadeemiline koostöö teiste ülikoolidega



- Teadusliku tegevuse ja õppetöö ajakohasuse ja kõrge taseme tagab avatus ja aktiivsus suhtluses koostööpartneritega välismaal
- Alates 90-ndate aastate algusest on Tartu statistikud kuulunud rahvusvahelistesse organisatsioonidesse, hetkel on MSI-s 4 ISI valitud liiget.
- 90ndate aastate alguses oli rahvusvahelise suhtluse osas majanduslike võimaluste piiratuse tõttu peamine väliskülaliste vastuvõtt.
- Peatselt leidsid nii õppejõud kui ka tudengid mitmesuguste projektide toel reisimisvõimalusi ja olukord tasakaalustus.
- Külalislektoritena välisülikoolides hakkasid aktiivselt tegutsema ka MSI õppejõud.



Külastatud MSIs loengul

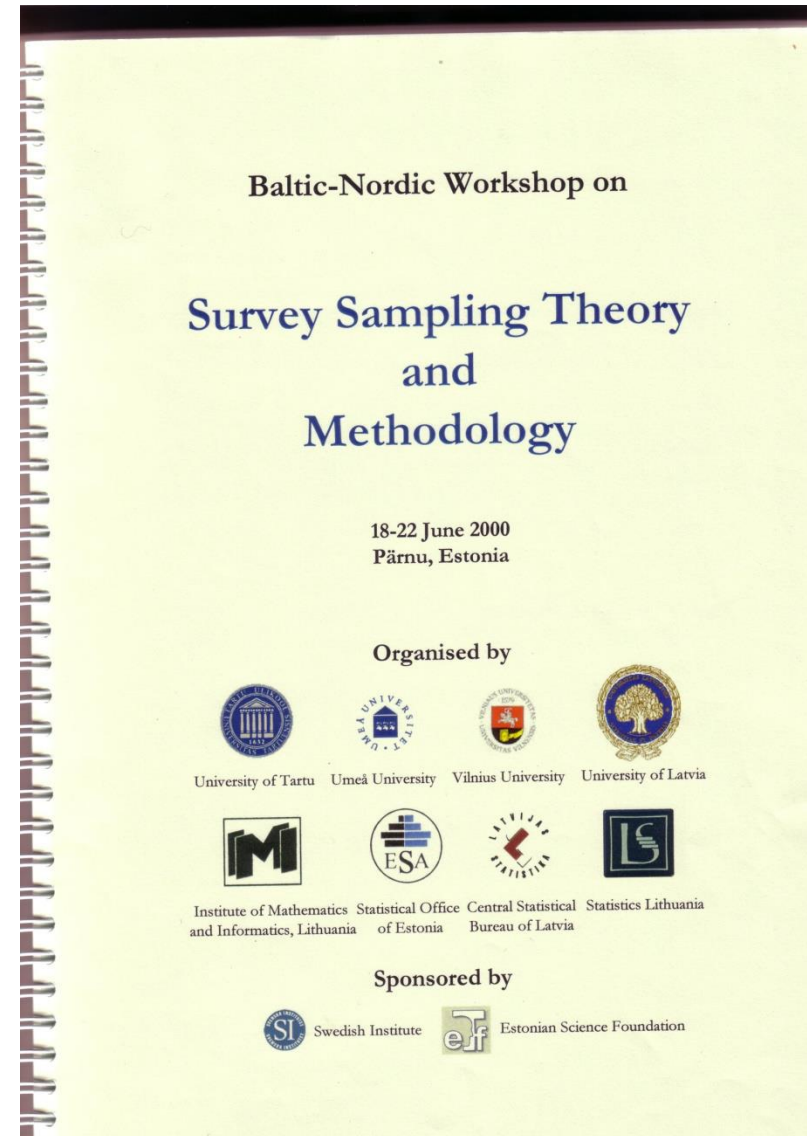


- Aastail 1992—96 stažeerisid Tartu tudengid Tempuse raames Barcelonas;
- Peaaegu igal aastal õpib mõni MSI tudeng Belgias magistriõppes.
- Tihe koostöö on olnud Rootsi (Uppsala ja Umea) ja Soome (Helsinki) ülikoolidega.
 - Imbi Traat on Helsinki ülikooli dotsent;
 - Ene-Margit Tiit on Helsinki ülikooli audoktor
- Välismaal töötab praegu vähemalt 5 doktori-kraadiga MSI kasvandikku (Krista Fischer, Tatjana Nahtmann, Birgit Strikholm, Jan Liinev, Artur Sepp). On neid, kelle sidemed Eestiga on juba nõrgaks jäänud, on ka neid, kes kavatsevad peagi tagasi pöörduda.



MSI on dotsent Imbi Traadi
juhtimisel osalenud üle
kümne aasta valikuuringute
Balti ja Põhjamaade
võrgustikus (*Baltic-Nordic
Network in Survey Sampling
Theory and Methodology*).

Võrgustiku iga-aastastel
seminaridel saavad tudengid
kogemusi nii rahvusvahelisele
kuulajaskonnale esinemisel
kui ka tulemuste
publitseerimisel.



Rahvusvaheliste Tartu Mitmemõõtmelise statistika konverentside seeria

- 1977, Kääriku (vene keel). 160 osavõtjat.
- 1981, Sangaste (vene keel). 190 osavõtjat.
- 1985, Kääriku (vene keel). 170 osavõtjat.
- 1989, Kääriku (vene keel). 130 osavõtjat.
- 1994, Pühajärve (inglise keel). 80 osavõtjat.
- 1999, Tartu (inglise keel). 80 osavõtjat.
- 2003, Tartu (inglise keel). 80 osavõtjat.
- 2007, Tartu (inglise keel). 100 osavõtjat



Konverentsikülalised Käärikul 1989. Teises reas teine prof. Boris Gnedenko



Tartu konverents 1994. Esineb prof. Ingram Olkin



Konverentsikülalised 1994. aastal.



Professor Carles Cuadras (paremal) kingib legendaarsele statistikule C. R. Raole oma raamatu. Tartu, 1994



Tartu konverentside ettekannete kogumikud aastaist 1977– 2007



2007. aasta konverentsi võtme-
esinejad: prof. Barry Arnold
(seisab) ja prof.
Balakrishnan



Konverentsi kuulajaid 2007



Ettekannete vaheajal, 2007



Konverentsikülastajate vastuvõtt Ülikooli Muuseumis, 2007





Matemaatiline statistika Eesti riigis

- Teiste erialade esindajate nõustamine
- Eesti Statistikaselts
- Statistikute mure statistilise kirjaoskuse pärast
- Eesti riigi asutused vajavad statistikuid



- Peamiselt klientide ja koostööpartnerite huve silmas pidades avaldati aastail 1970—1990 seerias “Programme kõigile” 8 vihikut, mis käsitlevad statistika tarkvara kasutamise metoodikat ja tulemuste tõlgendamist.
- MSI-s loetakse statistilise nõustamise erikursust.
- Koostööpartnereid on paljudelt elualadelt




Eesti Statistika Selts arvudes

- Eesti Statistika Selts loodi 1991. aastal ja ta kuulub ISI rahvuslike statistika seltside ühendusse.
- Eesti Statistika Selts on seni pidanud 19 konverentsi mitmesugustel teemadel.
- Eesti Statistika Seltsil on ca 100 liiget.
- Eesti Statistika Seltsi president on prof. Kalev Pärna.
- Eesti Statistika Selts on publitseerinud 18 Teabevihikut.
- Järgmise, 2008. aastal toimuva konverentsi teemaks on kavakohaselt “Kvaliteet statistikas ja statistika kvaliteet”.







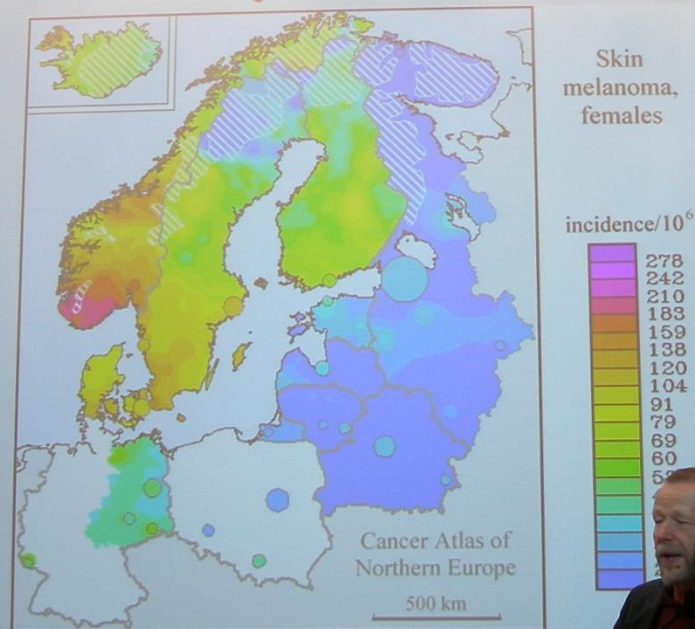
Statistiline kirjaoskus – mis see on?

Ene-Margit Tiit

ESS 19. konverents, Tallinn, 25—26 jaanuar 2007



Kõrge sotsiaalse seisundiga seotud vähk (nahamelanoom)



Allikas: Pukkala E et al. Cancer, 2001.





Statistikute meeskond



- MSI suurimaks väärtuseks on tegus ja võimekas meeskond.
- Seda meeskonda, mis koosneb erinevate huvidega erinevas vanuses inimesi, seob meeskonnatunne.
- Meeskonnavaim on kujunenud ühises vastutuses ja ühistes vaba aja ettevõtmistes ning traditsioonides



Igal aastal juunis traditsiooniline värskete lõpetajate pidu: MSI vanem põlvkond rõõmustab koos noorte lõpetajatega



Tartu Konverentsi organiseerimistöös on oluline roll kanda noortel (Anu Roos, Ants Kaasik, Anne Selart ja Helle Kilgi) konverentsikülalisi registreerimas, 2007.



Pihla päevad

- Suvel toimuvad Saaremaal Tõnu ja Imbi suvekodus traditsioonilised Pihla päevad, mis ühendavad klassikalist rahvusvahelist seminari mõnusa pereüritusega.
- Osa võtma on kutsutud kõik instituudi töötajad ja kraadiõppurid



Ettekandega esineb dots. Ebu Tamm (TTÜ)



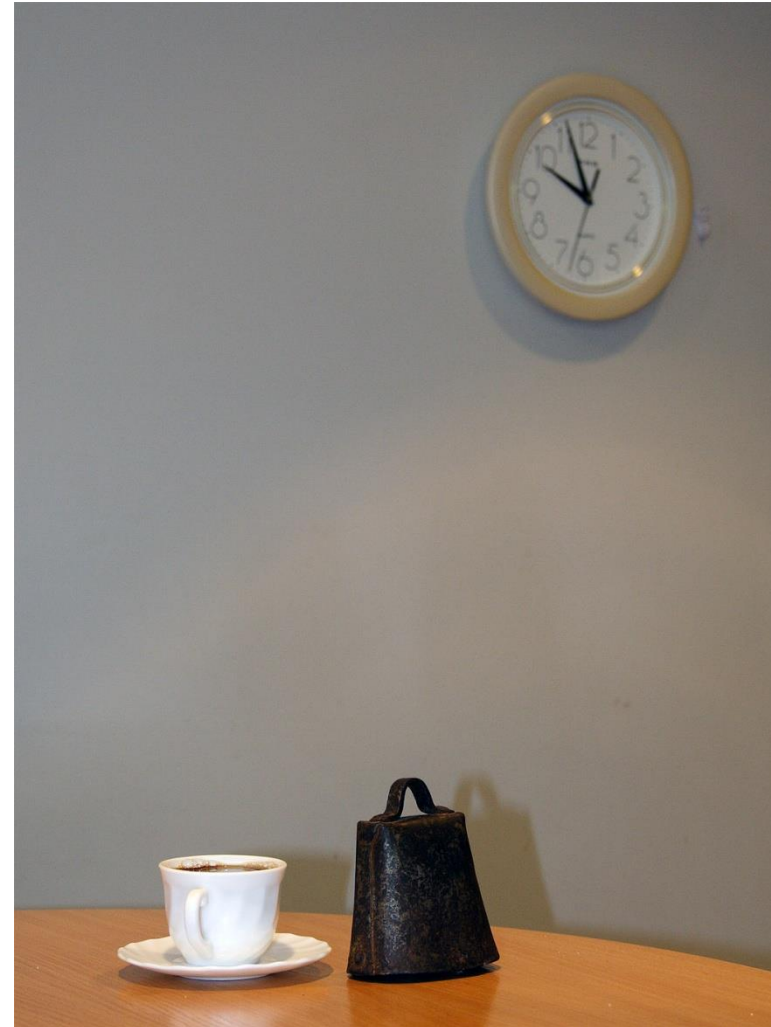


Kamina- või lõkkeseminar

- on üks statistikute vanemaid traditsioone. E.-M. Tiidu koduaias lõkketule juures on tähistatud paljude väliskülaliste Tartus viibimist



- MSI pere kohtub igal tööpäeval kell 10 ja kell 15 kohvitasside juures.
- Kohvitundi meenutab iidne karjakell



Statistika ja andmeteadus

MODERN DATA SCIENTIST

Data Scientist, the sexiest job of the 21st century, requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

MATH & STATISTICS

- ☆ Machine learning
- ☆ Statistical modeling
- ☆ Experiment design
- ☆ Bayesian inference
- ☆ Supervised learning: decision trees, random forests, logistic regression
- ☆ Unsupervised learning: clustering, dimensionality reduction
- ☆ Optimization: gradient descent and variants

DOMAIN KNOWLEDGE & SOFT SKILLS

- ☆ Passionate about the business
- ☆ Cares about data
- ☆ Influence without authority
- ☆ Hacker mindset
- ☆ Problem solver
- ☆ Strategic, proactive, creative, innovative and collaborative

PROGRAMMING & DATABASE

- ☆ Computer science fundamentals
- ☆ Scripting language e.g. Python
- ☆ Statistical computing packages, e.g. R
- ☆ Databases: SQL and NoSQL
- ☆ Relational algebra
- ☆ Parallel databases and parallel query processing
- ☆ MapReduce concepts
- ☆ Hadoop and Hadoop Pig
- ☆ Custom reducers
- ☆ Experience with cloud like AWS

COMMUNICATION & VISUALIZATION

- ☆ Able to engage with senior management
- ☆ Story telling skills
- ☆ Translate data driven insights into decisions and actions
- ☆ Visual art design
- ☆ R packages like ggplot or lattice
- ☆ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau



Andmeteadus

- Andmeteadus on rakendusstatistika ja andmeanalüüsi edasiarendus, mis kasutab aktiivselt arvutustehnika võimalusi, rakendab vajadusel ka teiste matemaatika valdkondade metoodikat ja on tihedasti seotud konkreetsete rakendusvaldkondadega (nt bioloogia, geneetika, sotsiaalteadused).

Täna tähelepanu eest!