

# Úrtak sem hlutfall af þýði

## Fyrirlestur í Aðferðafræði II

© 1998 Guðmundur Arnkelsson

All rights reserved. Copying or distribution prohibited without explicit permission. Students in Methodology II at the University of Iceland may print a copy for their own private use.

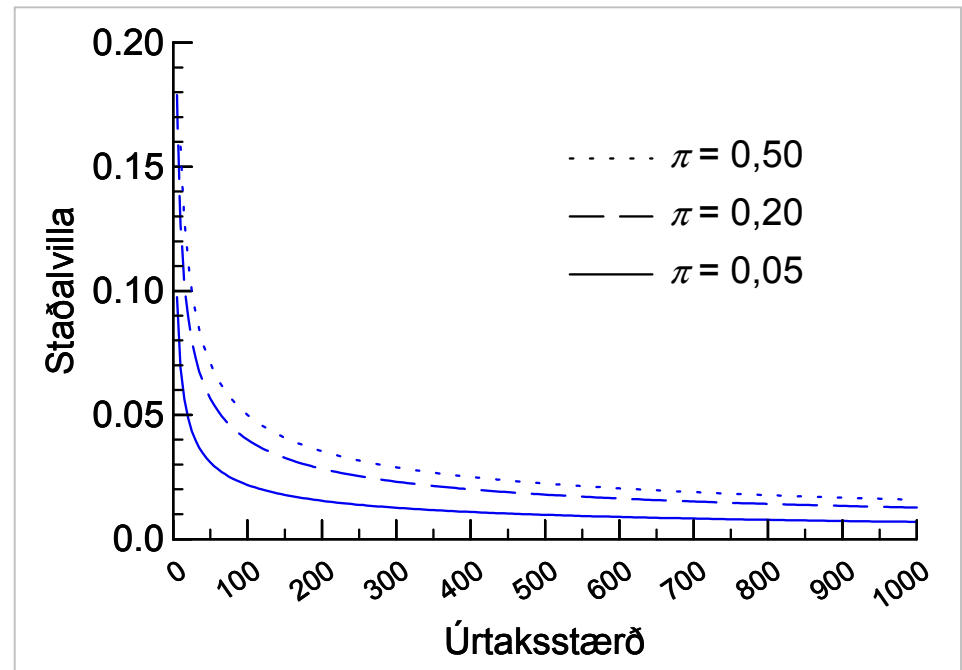
# Staðalvilla hlutfalls

Staðalvilla minnkar með aukinni stærð úrtaks og eykst eftir því sem hlutfallið í þýði er nær 0,5.

Staðalvillan minnkar ört frá  $N=1$  til  $N=200$  en mun hægar eftir það.

Þetta miðast við að úrtak sé dregið með skilum.

$$\hat{\sigma}_{\hat{\pi}} = \sqrt{\frac{\hat{\pi} \cdot (1 - \hat{\pi})}{N}}$$



# Rannsókn á sálfræðinimum

Nemar í námsráðgjöf könnuðu afdrif nemenda sem útskrifuðust með B.A. próf í sálarfræði á árunum 1980–1992.

Við upphaf rannsóknar þarf að velja úrtaksstærð með hliðsjón af þeirri nákvæmni (staðalvillu) sem óskað er eftir.

Af eðlilegum ástæðum er ekki hægt að draga úrtak með skilum. Samtals útskrifuðust 319 einstaklingar á þessu árabili og því gæti sami maðurinn lent mörgum sinnum jafnvel í meðalstórt úrtak.

Við höfum ástæðu til að ætla að um 60% nemanna muni vinna hjá hinu opinbera.

Við könnum staðalvillu hlutfallsins bæði með okkar hefðbundnu formúlum en einnig með Monte-Carlo tilraunum. Þannig má meta áhrif þess að úrtakið er ekki valið með skilum.

# Úrtak án skila

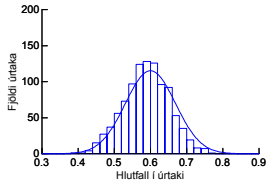
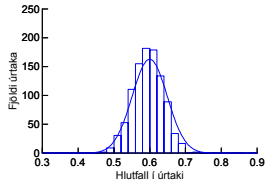
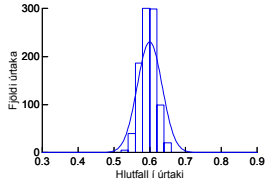
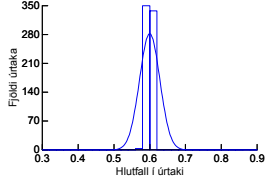
Óleiðrétt staðalvilla

Staðalvilla skv.  
Monte-Carlo tilraun

Gott samræmi er milli útreikninga og Monte-Carlo tilraunar þegar úrtaksstærð er lítil.

Eftir því sem stærð úrtaks nálgast stærð þýðis fer að muna meiru á útreikningum og niðurstöðum Monte-Carlo tilrauna.

Þegar úrtakið er orðið næstum jafnstórt þýðinu sýnir Monte-Carlo rannsókn fjórfalt lægri staðalvillu en fæst með formúlunni.

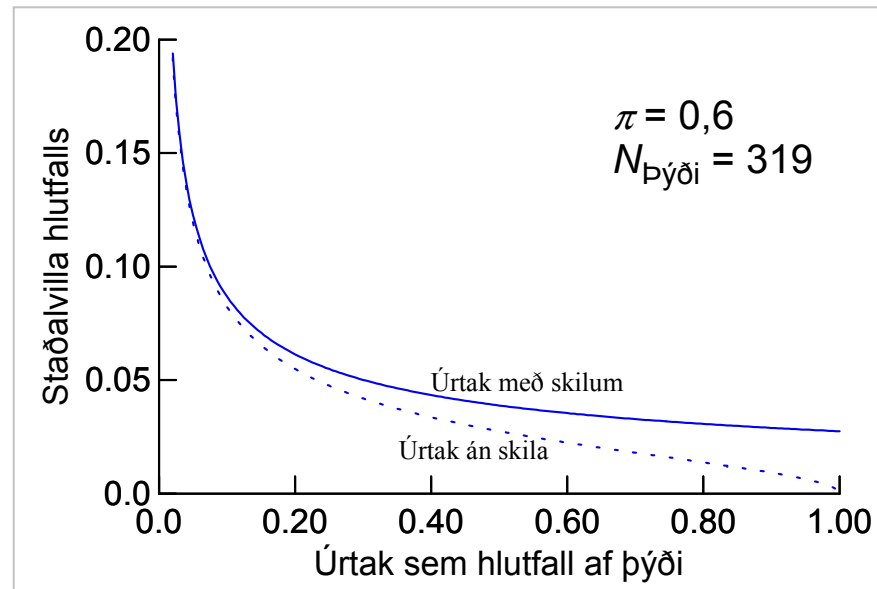
N	$\sigma_{\hat{\pi}}$	$s_p$	
50	0,069	0,063	
100	0,049	0,042	
200	0,035	0,021	
300	0,028	0,007	

# Leiðrétt staðalvilla

Myndin sýnir muninn á staðalvillu eftir því hvort úrtak er með skilum eða ekki.

Eftir að úrtakið hefur náð 200 manna stærð (um 2/3 af þýði) minnkar staðalvillan lítið ef úrtak er dregið með skilum.

Staðalvilla minnkar hins vegar hratt ef úrtak er án skila og nálgast 0,0 þegar úrtakstærð nálgast stærð þýðisins.



# Formúla fyrir úrtak án skila

Þegar úrtak er án skila þarf að leiðrétta útreikninga staðalvillu og taka þannig tillit til hve stórt úrtakið er af þýðinu öllu.

Nýju formúlurnar gefa réttari niðurstöður en þær gömlu sbr. niðurstöður Monte-Carlo tilraunanna.

Staðalvilla  
hlutfalls

$$\hat{\sigma}_{\hat{\pi}} = \sqrt{\frac{\hat{\pi} \cdot (1 - \hat{\pi})}{n_{\text{Úrtak}}} \cdot \left(1 - \frac{n_{\text{Úrtak}}}{N_{\text{Þýði}}}\right)}$$

Staðalvilla  
meðaltals

$$\hat{\sigma}_{\bar{X}} = \sqrt{\frac{\hat{\sigma}^2}{n_{\text{Úrtak}}} \cdot \left(1 - \frac{n_{\text{Úrtak}}}{N_{\text{Þýði}}}\right)}$$

# Heimasíður kennara

Ég vil kanna hve margir kennara í félagsvísindadeild eru með heimasíður fyrir námskeið sín.

Fyrirfram giska ég á að þetta séu 4–5 fastir kennarar eða um 15% af 32 föstum kennurum við deildina.

Ég vil meta hve stórt úrtak ég þarf til að meta þetta hlutfall nákvæmlega. Ég byrja á að athuga staðalvilluna ef ég vel helming kennarana í úrtakið.

$$\begin{aligned}\hat{\sigma}_{\hat{\pi}} &= \sqrt{\frac{\hat{\pi} \cdot (1 - \hat{\pi})}{n_{\text{Úrtak}}} \cdot \left(1 - \frac{n_{\text{Úrtak}}}{N_{\text{Þýði}}}\right)} \\ &= \sqrt{\frac{0,15 \cdot (0,85)}{16} \cdot \left(1 - \frac{16}{32}\right)} \\ &= \sqrt{\frac{0,1275}{16} \cdot 0,5} = \sqrt{\frac{0,06375}{16}} \\ &= \sqrt{0,00398} = 0,0631\end{aligned}$$

Staðalvillan er tiltölulega há. Ég þarf því stærra úrtak til að fá næga nákvæmni.

# Staðalvilla eftir hlutfalli í þýði

Fullnægjandi nákvæmni næst ekki fyrir en við erum komin með um 28 kennara í úrtakið.

Best væri því að spyrja alla fasta kennara í félagsvísindadeild, kanna þýðið allt.

Taflan ber greinilega með sér að formúlan sem miðast við skil vanmetur stórlega nákvæmni niðurstaðna í okkar tilfelli.

n	Hlutfall af þýði	Staðalvilla	
		Með skilum	Án skila
4	0,125	0,179	0,167
8	0,250	0,126	0,109
16	0,500	0,089	0,063
24	0,750	0,073	0,036
28	0,875	0,067	0,024
32	1,000	0,063	0,000



# Laun sálfræðinema eftir útskrift

Í könnuninni á sálfræðinimum var spurt um grunnlaun þeirra. Þau reyndust vera að meðaltali kr. 121.000 með staðalfrávikið 45.300 í úrtaki 71 útskrifaðs sálfræðinema.

Ef við endurtökum rannsóknina og öflum nákvæmari upplýsinga um laun að loknu B.A. prófi, hversu stórt úrtak ættum við að nota?

$$\hat{\sigma}_{\bar{X}} = \sqrt{\frac{\hat{\sigma}^2}{n_{\text{Úrtak}}} \left(1 - \frac{n_{\text{Úrtak}}}{N_{\text{Þýði}}}\right)}$$

n	Hlutfall af þýði	Staðalvilla	
		Með skilum	Án skila
40	0,125	7.163	6.698
71	0,223	5.376	4.740
160	0,502	3.581	2.528
239	0,749	2.930	1.467
319	1,000	2.536	0