t-próf á meðaltöl

Fyrirlestur í Tölfræði II (SÁL203G)

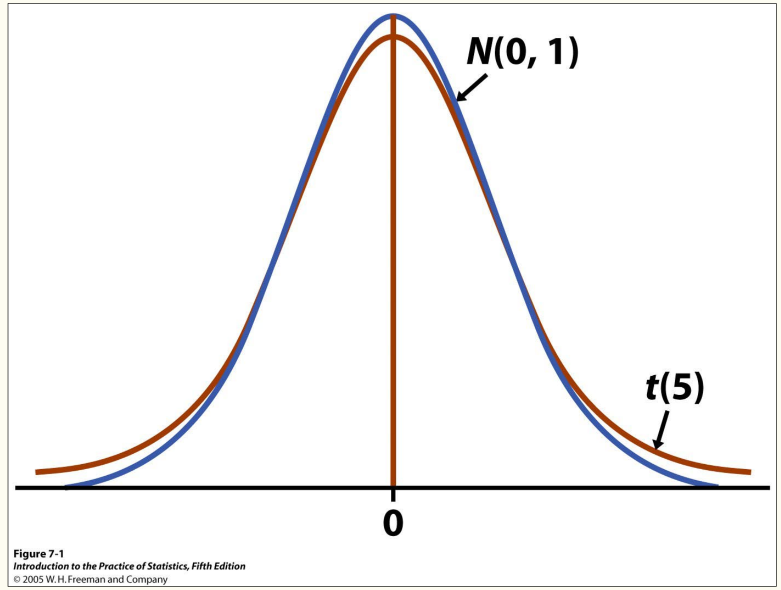
2011-01-31

t-dreifing

Þegar staðalfrávik þýðis er óþekkt, gefur staðalfrávik úrtaksins bestu spá. Staðalfrávik meðaltala er þá óþekkt en metið með staðalvillu meðaltala.

Staðalfrávik er breytilegt milli úrtaka og því fáum við aðeins ónákvæmt mat á staðalvillunni. Við bregðumst við því með því að nota *t*-dreifingu í stað normaldreifingar til að lýsa dreifingu meðaltala.

t-dreifing er með lengri hala en normaldreifing og tekur þannig tillit til óvissunnar í matinu.



Lengri halar þýðir að meira frávik þarf frá H_0 til að hægt sé að hafna henni. Ef úrtak er stórt er *t*-dreifing þó eins og normaldreifing.

Samanburður *t*- og normaldreifingar

t-dreifing fer eftir frígráðunum en þær miðast við úrtaksstærð. Við eitt meðaltal er $df = n – 1$.

Taflan sýnir vendigildi *t*-dreifingar—t^*—fyrir ólíkar frígráður. Þau eru mun hærri en fyrir normaldreifingu þar til $df \approx 30$.

Dreifingarnar nálgast hvor aðra eftir því sem frígráðum *t*-dreifingar fjölgar og eru orðnar nánast alveg eins þegar frígráðurnar eru komnar yfir hundrað.

TABLE D *t* distribution critical values

	Upper tail probability *p*											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
z^*	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level *C*											

t-próf

t-próf er í framkvæmd nánast eins og z-próf. Munurinn er sá að við notum spágildi staðalfráviks og metum staðalvilluna í stað þess að geta reiknað breidd úrtakadreifingarinnar (*sampling distribution*) nákvæmlega.

Niðurstöðunni er flett upp í t-töflu með hliðsjón af frígráðunum. Ef líkurnar á þetta miklu eða meira frávik frá H_0 er jafnt og eða minna en α, getum við hafnað núlltilgátunni og tekið upp aðaltilgátuna.

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

$$df = n - 1$$

$$\mu_0 : \text{Viðmiðsgildi}$$

Staðalfrávik úrtaksins, s, er annað hvort of- eða vanmat á staðalfráviki þýðis, σ. Við því er brugðist með því að bera niðurstöðuna, t-ið, við vendigildi í t-töflu. Frávikið frá normaldreifingu minnkar með hækkandi frígráðum.

Viðmiðsgildið μ_0 stendur fyrir þýðismeðaltalið sem núlltilgátan tilgreinir.

Öryggisbil

Öryggisbil eru reiknuð á sama hátt og fyrir *z*-próf.

Hér þarf að gæta þess að þar sem þýðisstaðalfrávikið er óþekkt, þarf að fletta upp réttu vendigildi—*t**—miðað við frígráður og α í *t*-töflu.

Algengast er að nota 95% og stundum 99% öryggisbil.

Moore og McCabe nota gjarnan 90% öryggisbil en það er tiltölulega sjaldséð.

$$\bar{X} - t^* \cdot \frac{s}{\sqrt{n}} < \mu < \bar{X} + t^* \cdot \frac{s}{\sqrt{n}}$$

eða

$$\bar{X} \pm t^* \cdot \frac{s}{\sqrt{n}}$$

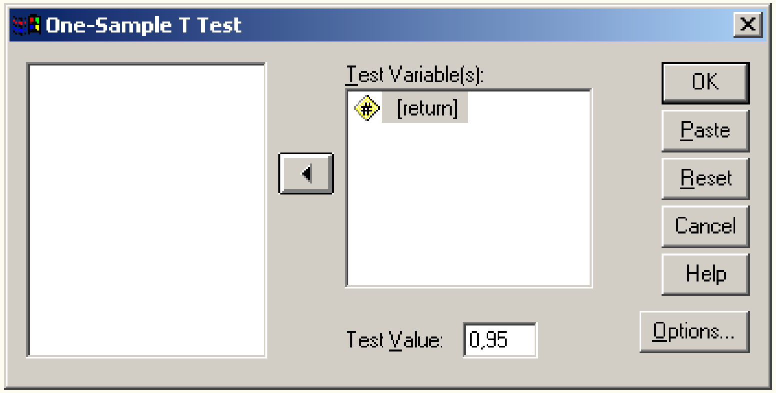
Vegna óvissu um staðalfrávik þýðis verður samsvarandi óvissa um staðalvillu meðaltal, dreifingu meðaltalanna. Því er vendigildi flett upp í *t*-töflu sem gefur hærri gildi en normaltafla og því *breikkar* öryggisbilið.

t-próf í einum hópi í SPSS

Fyrst er skráin `ta07_001.por` lesin inn. Síðan skoðum við normal- og kassarit með því að fara í `Analyze/ Descriptive Statistics/ Explore....`

t-próf í einum hópi er framkvæmt með því að fara í `Analyze/ Compare Means/ One Sample T-Test....`

Þar færi ég breytuna í textareitinn, gef upp viðmiðsgildið og smelli síðan á `OK`.



Er ávöxtunin önnur en hjá S&P?

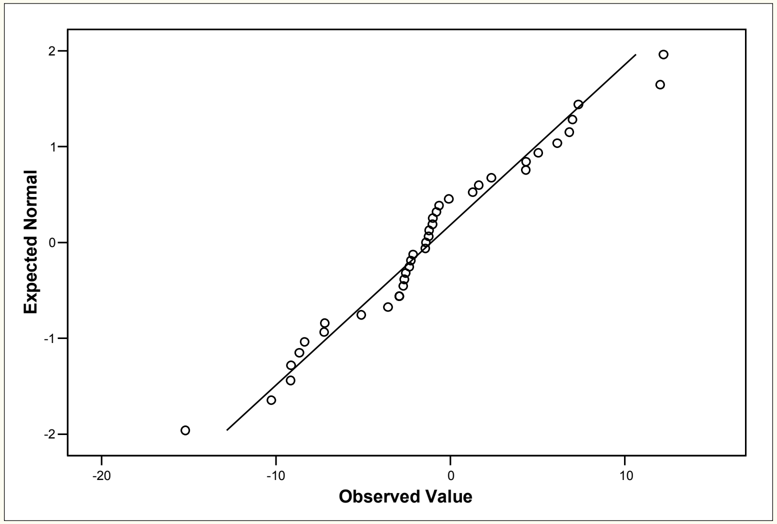
Við viljum vita hvort ávöxtun eins hlutabréfasafns sé önnur en gengur og gerist á markaðinum.

Við höfum ávöxtun síðustu 39 mánaða en ávöxtun Standards og Poors var að meðaltali 0,95% fyrir sama tímabil.

Fyrst skoðum við lýsandi niðurstöður. Meðaltal og staðalfrávik eru eðlileg. Hæstu og lægstu gildi eru óvenjuleg en í samræmi við töflu í kennslubók. Dreifingin er samhverf, með þrjú jaðargildi en engin alvarleg frávik.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
	39	-15,20	12,22	-1,0985	5,98779
Valid N (listwise)	39				



Niðurstöður *t*-prófs í einum hópi

t-prófið prófar þá tilgátu að ávöxtunin sé annað hvort meiri eða minni en 0,95% á mánuði.

Prófið er marktækt sem gerir okkur kleift að hafna núlltilgátunni. Við getum því fullyrt að ávöxtunin hafi ekki verið 0,95%. Það gæti verið rangt og vitum ekki líkur þess.

Við vitum að *ef* H_0 er rétt, þá eru aðeins tæplega 4% líkur á svona miklu fráviki frá henni. Niðurstaðan veikir því trú okkar á núlltilgátunni.

Við vitum *ekki* líkurnar á þessu fráviki þegar H_0 er rangt.

One-Sample Test

	Test Value = 0.95					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
	-2,136	38	,039	-2,04846	-3,9895	-,1074

$$H_a : \mu \neq 0{,}95 \qquad H_0 : \mu = 0{,}95$$

Marktektarprófið segir okkur hvort við getum hafnað núlltilgátunni. Öryggisbilið gefur upp sennilegt frávik frá viðmiðsgildinu. Það gæti verið frá því að vera mjög lítið (–0,1%) upp í að vera mjög umtalsvert (–4,0%).

$$\text{Cohen's d} : d = \frac{-2{,}0485}{5{,}9878} = -0{,}34$$

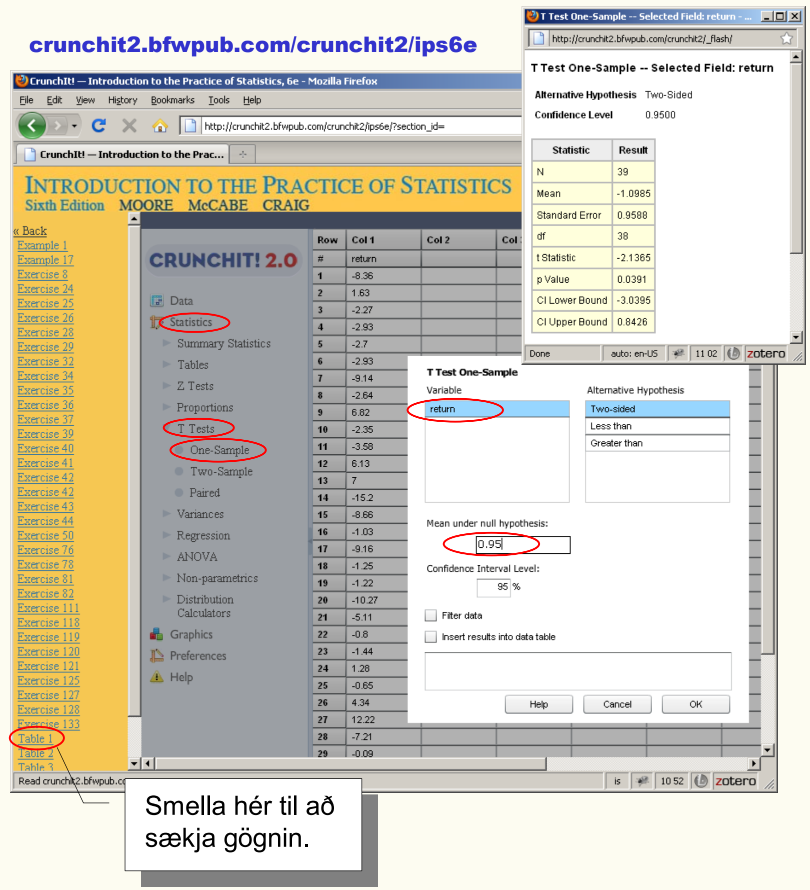
t-próf í einum hópi í CrunchIt

Við opnum CrunchIt, veljum kafla 7 í dálki lengst til vinstri og svo Table 1.

Svo veljum við `Statistics / T Tests / One-Sample`. Þar veljum við breytuna, gefum upp viðmiðsgildið og smellum á `OK`.

Niðurstöðurnar birtast í sérglugga, með niðurstöðu prófsins, frígráðum, *p*-gildi og öryggisbil.

Öryggisbilið er í kringum *meðaltalið* en ekki fráviki þess frá μ_0 eins og í SPSS.



Parað *t*-próf

Stundum eru gerðar tvær mælingar á sama einstaklingi, t.d. mælingar fyrir og eftir inngrip. Þá er tilgátan oftast sú að breyting á niðurstöðu mælinga sé samfara inngripinu.

Þátttakendur koma stundum í pörum, t.d. hjón, kærustupör, feðgin, o.s.frv. Þá verður að taka tillit til tengslanna.

Í sumum tilvikum eru tveir og tveir þátttakendur paraðir saman svo að þeir séu sem líkastir á einhverjum mikilvægum breytum.

Við höfum upplýsingar um tíðni erfiðrar hegðunar hjá heilabiluðum einstaklingum. Við viljum vita hvort það skipti máli hvort það sé fullt tungl eða ekki. Við berum því hegðunina þá þrjá daga sem eru næstir fullu tungli saman við hegðun hina 26 daganna.

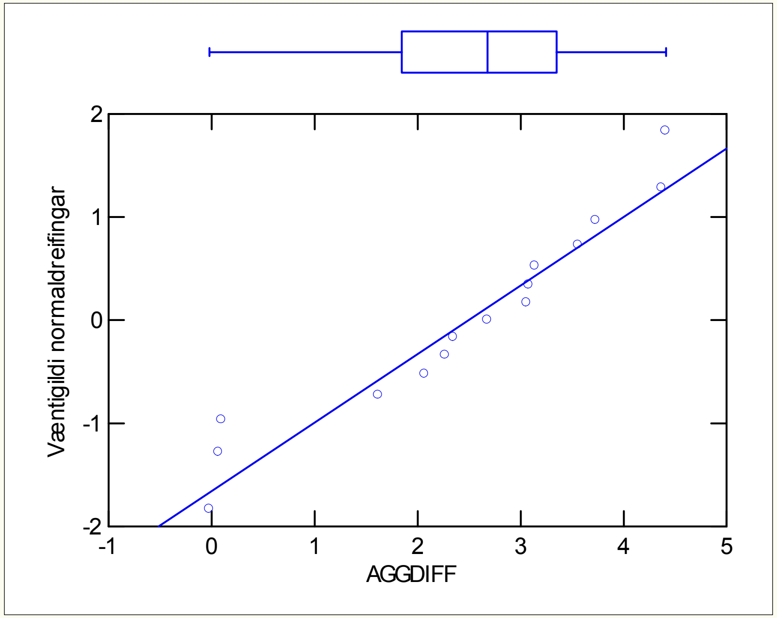
Hér eru tvær mælingar (meðaltöl hegðunar) fyrir hvern einstakling og því eru mælingarnar paraðar. Við notum því parað *t*-próf til að prófa tilgátuna.

Parað *t*-próf sem próf í einum hópi

Við viljum athuga hvort *mismunurinn* á mælingunum tveimur víki frá 0,0 eða öðru viðmiðsgildi. Parað *t*-próf er því í reynd eins hóps próf á mismun mælinganna tveggja.

Við viljum því skoða mismunatöluna og eiginleika hennar. Við skoðum því kassarit eða normalrit af mismuninum og könnum hvort þar sé eitthvað að sjá.

Það getur einnig verið gagnlegt að skoða hvora mælingu fyrir sig, þótt það skipti ekki máli fyrir forsendur prófsins.



Neikvæð skekkja með þremur jaðargildum en engin stór frávik frá normaldreifingu.

Parað *t*-próf í SPSS

Ég les `ta07_002.por` inn. Parað *t*-próf er framkvæmt með því að fara í `Analyze/ Compare Means/ Paired-Samples T-Test`… og setja báðar breyturnar inn. Ég hefði einnig getað notað `/ One Sample T-Test`… og valið breytuna `AggDiff`.

Skoðaðu vel meðaltöl, staðalfrávik, fjölda og fylgnina til að vera viss um að allt sé með felldu. Allt sem er óvænt eða undarlegt þarf að skoða nánar.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1		3,0220	15	1,49877	,38698
		,5893	15	,44490	,11487

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	&	15	,234	,402

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	-	2,43267	1,46032	,37705	1,62397	3,24137	6,452	14	,000

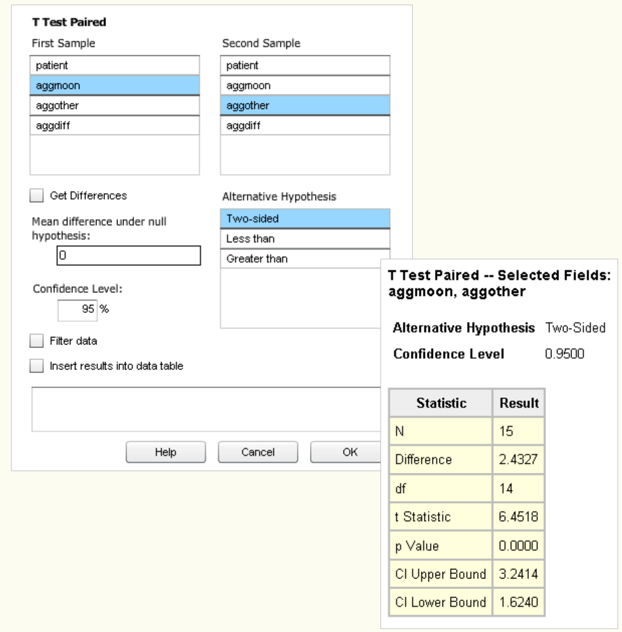
Tíðni erfiðrar hegðunar reyndist hærri (*M*= 3,0) þá daga sem voru næstir fullu tungli en aðra daga (*M*= 0,6), *t*(14)= 6,5, *p*< 0,001 (Cohens *d*= 1,7). Tíðni erfiðrar hegðunar er á bilinu 1,6 til 3,2 hærri á fullu tungli miðað við 95% öryggi.

Parað *t*-próf í CrunchIt

Ég sæki töflu 2 í kafla 7. *t*-próf í einum hópi er framkvæmt með því að fara í `Statistics/T Test/ Paired` og velja `AggMoon` og `AggOther` hvort í sínum vallista. Ég hefði einnig getað notað `One-Sample` með breytuna `AggDiff`.

Prófið er marktækt og því má álykta að hegðun sé erfiðari við fullt tungl.

Mikilvægustu upplýsingarnar eru gefnar upp nema staðalfrávik mismunar sem þarf að reikna út: `Statistics/Summary Statistics/Column`



Statistic	Result
N	15
Difference	2.4327
df	14
t Statistic	6.4518
p Value	0.0000
CI Upper Bound	3.2414
CI Lower Bound	1.6240

$$Cohens\ d = \frac{2{,}433}{1{,}460} \approx 1{,}7$$

Túlkun marktektar og öryggisbils

Ef niðurstaðan er marktæk, getum við staðhæft að núlltilgátan sé röng. Hún tilgreinir að ekkert frávik sé í þýði frá viðmiðsgildinu (sem oftast er 0,0).

Jafnvel þótt núlltilgátan *sé* röng, þýðir það ekki að frávikið skipti máli. Við þurfum því að fá eitthvert mat á því hversu stórt frávikið er.

Marktekt er tvíkostaákvörðun en oftast er óvissa um niðurstöðuna.

Öryggisbil gefa frekari upplýsingar um þýðingu niðurstaðnanna.

Taflan miðast við að *t*-próf sé *marktækt* og öryggisbil sé reiknað í framhaldi af því.

Neðri mörk	Efri mörk	Ályktun
Óverulegt frávik	Óverulegt frávik	Það *er* munur sem skiptir engu máli
Óverulegt frávik	Umtalsvert frávik	Það er munur sem *gæti* skipt máli.
Umtalsvert frávik	Umtalsvert frávik	Það er munur sem skiptir verulegu máli.

www.graphpad.com/articles/interpret/Analyzing_two_groups/paired_t.htm

Túlkun öryggisbils ef ekki marktekt

Ef prófið er ómarktækt, get ég *ekki* hafnað núlltilgátunni. Hún gæti samt sem áður verið röng. Hugsanlega var prófið ekki nógu afkastamikið eða frávikið lítið en samt fyrir hendi.

Öryggisbilið gefur meiri upplýsingar heldur en tvíkostaákvörðunin sem felst í marktektarprófinu. Með því að skoða efri og neðri mörk bilsins er hægt að álykta um sennilega stærð fráviksins—ef það þá er til staðar—og fá þannig fyllri upplýsingar um þýðismeðaltalið.

Taflan miðast við að *t*-próf sé *ekki* marktækt og öryggisbil sé reiknað í framhaldi af því.

Neðri mörk	Efri mörk	Ályktun
Óverulegt frávik	Óverulegt frávik	Það er enginn eða a.m.k. óverulegur munur.
Óverulegt frávik	Umtalsvert frávik	Það *gæti* verið munur eða ekki.
Umtalsvert frávik	Óverulegt frávik	Það *gæti* verið munur *niður á við* eða jafnvel ekki.
Umtalsvert frávik	Umtalsvert frávik	Kannski er munur, kannski ekki.

www.graphpad.com/articles/interpret/Analyzing_two_groups/paired_t.htm

Forsendur *t*-prófs

Marktektarpróf byggja á ákveðnum forsendum. Fyrir *t*-próf er forsendan um normaldreifingu mikilvægust. Í henni felst í meginatriðum að dreifingin sé hvorki skekkt né með frávillingum.

t-próf er álitið traust (*robust*) gagnvart skekkju en ótraust gagnvart frávillingum. Traustleikinn eykst með úrtaksstærð.

Frávik eru skoðuð myndrænt með normal- eða kassariti.

Úrtaksstærð	Úrræði
$n < 15$	Erfitt að meta forsendur. Reyna umbreytingar eða aðrar aðferðir ef veruleg frávik eru til staðar.
$15 \leq n < 40$	Nota *t*-próf nema skekkja sé mjög mikil eða frávillingar greinilegir.
$n \geq 40$	Hægt að nota *t*-próf þrátt fyrir mikla skekkju eða staka frávillingar.

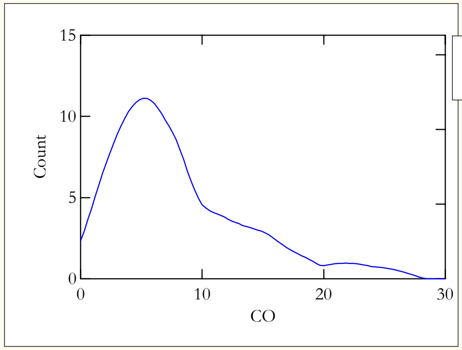
Áhrif skekkju á niðurstöður *t*-prófs

Myndin sýnir dreifingu kolsýrings (CO) í mengunarmælingum bíla.

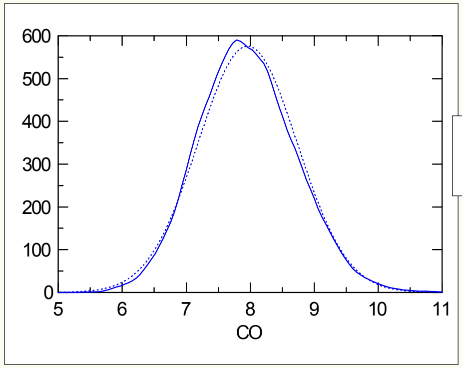
Þrátt fyrir að mælingarnar séu jákvætt skekktar í úrtakinu, þá verður úrtakadreifing meðaltalsins nær normaldreifð.

Úrtakadreifingin fylgir nánast alveg normaldreifingu, sbr. punktalínuna sem sýnir normalferil.

Myndin gefur til kynna að 95% öryggisbil nái yfir 6,5 til 9,5 í stað 6,4 til 9,5 eins og útreikningar gáfu til kynna.



Upprunaleg dreifing



Úrtakadreifing (10.000 endurúrtök; *bootstrap*)

t-próf í óháðum hópum

Þegar borin eru saman tvö meðaltöl eru staðalfrávikin einnig tvö. Þar sem þau koma úr hvort úr sínu þýðinu, er gert ráð fyrir að þýðisstaðalfrávikin séu ólík.

Niðurstaða t-prófsins fylgir ekki t-dreifingu nákvæmlega. Við því er brugðist með því að fletta niðurstöðu þess upp með færri frígráðum. Við það minnkar marktekt prófsins, sérstaklega í fámennum úrtökum.

Best er að láta forritum eftir nákvæma útreikninga.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

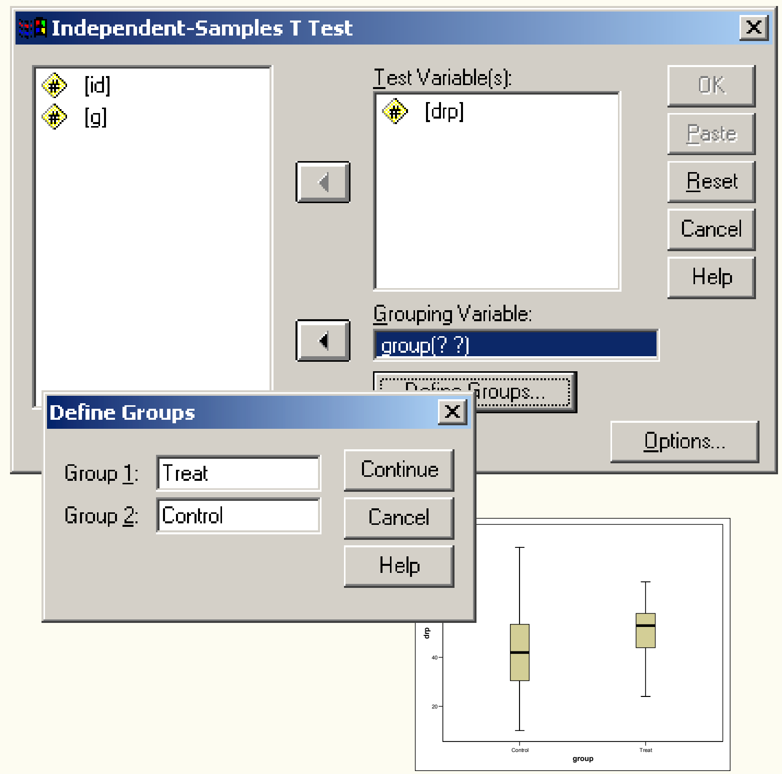
$$H_0: \mu_1 = \mu_2$$

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left(\frac{s_1^2}{n_1}\right)^2 \cdot \frac{1}{n_1} + \left(\frac{s_2^2}{n_2}\right)^2 \cdot \frac{1}{n_2}} - 2$$

Tveggja hópa *t*-próf í SPSS

Ég sæki töflu 4 í 7. kafla. *t*-próf í tveimur óháðum hópum er framkvæmt með `Analyze/ Compare Means/ Independent-Samples T-Test`.... Þar skilgreini ég hópana og smelli síðan á OK.

Kassaritið úr `Explore` gefur til kynna að allt sé með felldu og ekkert því til fyrirstöðu að reikna *t*-prófið. Normalrit gefa ekki heldur ástæðu til sérstakra aðgerða. Inngripið virðist hafa aukið færni og minnkað breytileika.



Niðurstöður tveggja hópa prófs

Hér prófum við hvort sérstök verkefni auki lestrargetu nemenda í 3. bekk.

Lýsandi tölfræði gefur til kynna að inngripshópurinn sé með meiri færni en samanburðarhópurinn en einnig minni breytileika.

Í niðurstöðutöflunni skoðum við línuna sem gerir ráð fyrir ólíkum breytileika eftir hópum. Við höfnum H_0 þar sem $p = 0{,}026/2 \approx 0{,}01$. Öryggisbilið gefur þó ekki ótvírætt til kynna að áhrifin séu umtalsverð.

$$H_0 : \mu_{treatment} = \mu_{control}$$
$$H_1 : \mu_{treatment} > \mu_{control}$$

Group Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Treat	21	51,48	11,007	2,402
Control	23	41,52	17,149	3,576

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal variances assumed	2,362	,132	2,267	42	,029	9,954	4,392	1,091	18,818
Equal variances not assumed			2,311	37,855	,026	9,954	4,308	1,233	18,676

Niðurstaða *t*-prófsins

Leiðréttar frígráður

Marktekt miðað við tvíhliða próf

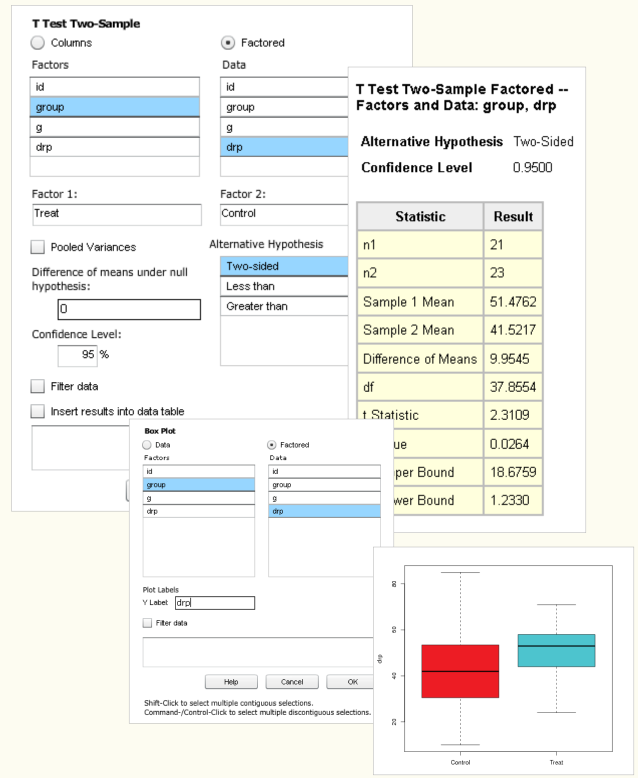
Öryggisbil

Tveggja hópa *t*-próf í CrunchIt

Ég les `ta07_004.txt` inn. *t*-próf í tveimur óháðum hópum er framkvæmt með `Statistics/ T Test/Two-Sample`. Þar vel ég `group` sem `Factors` og `dpr` sem `Data`.

Best er að gefa upp tvíhliða próf og deila síðan með tveimur í *p*-gildið til þess að fá hefðbundið öryggisbil. Ef einhliða próf er valið, verður öryggisbilið einhliða einnig.

Forsendur má meta með `Graphics/ Boxplot`, sbr. sýnishorn hér til hliðar.



Samlagðar (*pooled*) dreifitölur

Ef hægt er að gera ráð fyrir því að staðalfrávikin séu eins í þýði, fylgir *t*-prófið *t*-dreifingunni nákvæmlega.

Það er umdeilt hvort hægt sé að gefa sér jöfn staðalfrávik sem forsendu. Sumir vilja prófa forsenduna áður en prófið er framkvæmt, aðrir vilja að kassarit séu athuguð en enn aðrir—t.d. Moore & McCabe—kjósa að gera ráð fyrir ólíkum staðalfrávikum í flest öllum tilfellum.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Samlagðar dreifitölur í SPSS

Ef hægt er að gera ráð fyrir að staðalfrávik séu eins í þýði, er einfaldlega lesið úr efri línu töflunnar í SPSS.

Hægt er að nota niðurstöðu Prófs Levenes til að velja á milli lína, en það er þó umdeilt.

Að jafnaði er próf með samlögðum dreifitölum afkastameira heldur en þegar staðalfrávik eru ólík. Í þessu tiltekna tilviki er þessu þó öfugt farið.

Niðurstaða Prófs Levenes sem prófar $H_0{:}\sigma_1 = \sigma_2$

Marktekt fyrir Próf Levenes

Öryggisbil

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal variances assumed	2,362	,132	2,267	42	,029	9,954	4,392	1,091	18,818
~~Equal variances not assumed~~			~~2,311~~	~~37,855~~	~~,026~~	~~9,954~~	~~4,308~~	~~1,233~~	~~18,676~~

Niðurstaða *t*-prófsins

Frígráður

Marktekt miðað við tvíhliða próf

$$s_p = \sqrt{\frac{20 \cdot 11{,}007^2 + 22 \cdot 17{,}149^2}{42}}$$

$$= \sqrt{\frac{8893{,}21}{42}} = \sqrt{211{,}7386} \approx 14{,}55$$

$$\mathit{Cohens}\ d = \frac{9{,}954}{14{,}55} \approx 0{,}7$$

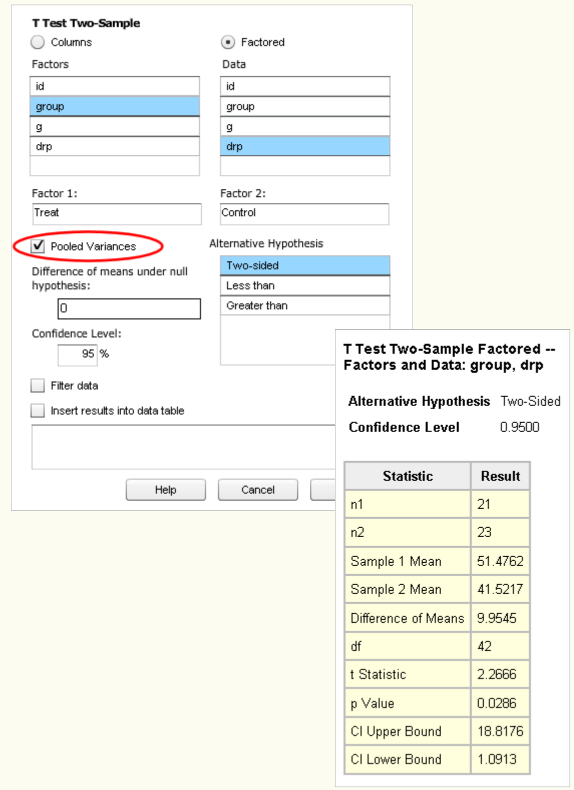
Samlagðar dreifitölur í CrunchIt

Eina breytingin í CrunchIt er að hakað er við `Pooled Variances`.

Niðurstöður eru svipaðar og með aðgreindum dreifitölum, en heldur lægri.

Í þessu tilviki er ekki ótvírætt hvora aðferð eigi að nota. Það er vísbending um misleitni í gögnunum en hún er þá lítil. Meðferð getur þó haft áhrif á dreifingu jafnt og meðaltal.

Moore og McCabe myndu velja aðgreindar dreifitölur ef það er minnsti vafi.



Statistic	Result
n1	21
n2	23
Sample 1 Mean	51.4762
Sample 2 Mean	41.5217
Difference of Means	9.9545
df	42
t Statistic	2.2666
p Value	0.0286
CI Upper Bound	18.8176
CI Lower Bound	1.0913

Dæmi um úrvinnslu

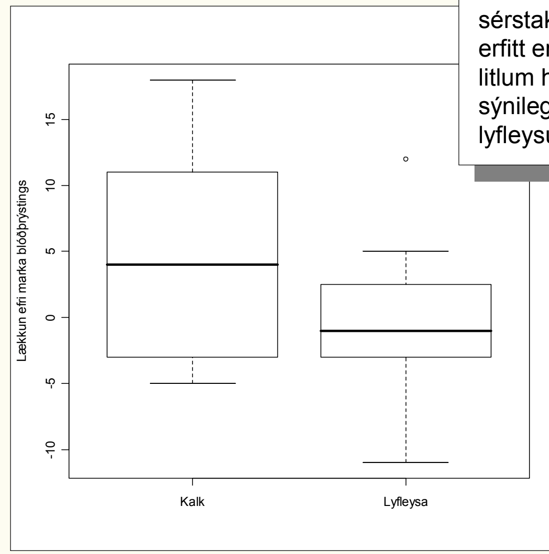
Myndin sýnir áhrif kalks á efri mörk blóðþrýstings.

Við metum þau frávik sem hafa áhrif á niðurstöður *t*-prófs, þ.e. skekkju, frávillinga, misleitni og ójafnan fjölda í hópum.

Myndin bendir til misleitni, próf Levenes virðist staðfesta það, $F(1, 19) = 4{,}3$, $p = 0{,}05$, en lítill munur er á staðalfrávikum. Eitt fráviksgildi sést.

Til öryggis, ólíkt M&M, miðum við því við ólík staðalfrávik í þýði.

Það er tiltölulega jafnt í hópum, einhver skekkja í báðum hópum en óveruleg sérstaklega miðað við hve erfitt er að meta hana í litlum hópum. Misleitni er sýnileg og eitt fráviksgildi í lyfleysuhópnum.



Þótt blóðþrýstingur lækkaði meira hjá þeim sem fengu kalk ($M = 5{,}0$) en hjá þeim sem fengu lyfleysu ($M = –0{,}3$) var ekki hægt að staðfesta þann mun, $t(19) = 1{,}6$, $p = 0{,}06$.