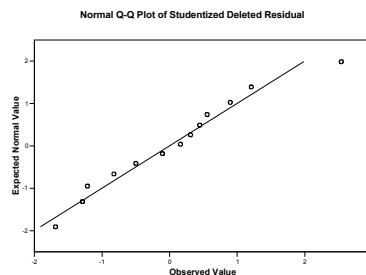


Eftirfarandi er yfirlit yfir spurningar sem notaðar hafa verið í prófum námskeiðinu.

Verkefni sem eru þess eðlis að þau myndu aðeins birtast í heimaprófi eru auðkennd með **H**. Almennt eru þetta þau verkefni sem krefjast aðgangs að tölfræðiforritum eða mikilla útreikninga.

1. Hvaða eiginleika metur summa kvaðrata í dreifgreiningartöflu?
2. Hvaða stærð metur meðalsumma kvaðrata leifar, MSE, í dreifgreiningartöflu?
3. Notaðu gagnasafnið ex10_013.sav. **H**
 - i. Breyttu hallanum fyrir Skakka turninn úr tíundu hlutum úr millimetra í sentímetra umfram 2,9 metra.
 - ii. Notaðu aðfallsgreiningu til að spá fyrir um halla turnsins fyrir árin 1990, 2010, 2030 og 2050. Gefðu upp viðeigandi öryggisbil. Túlkaðu niðurstöðurnar og tilgreindu þá heildarmynd sem niðurstöður og öryggisbil gefa.
4. Hver er munurinn á leif (*residual*) og villu (*error*)? Hvaða kröfur eru gerðar til villunnar í aðfallsgreiningu? Hvaða kröfur eru gerðar til leifarinnar?



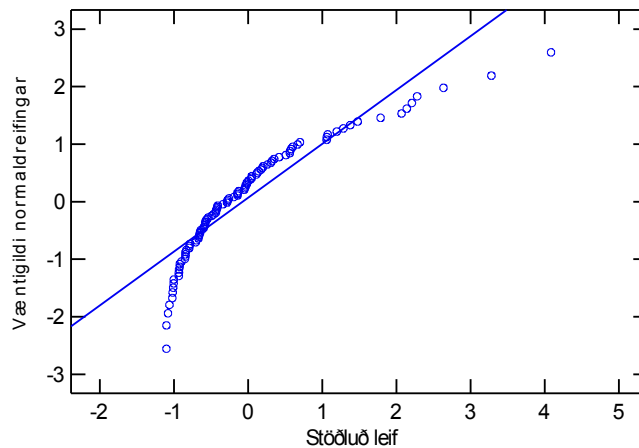
5. Hvaða frávík—ef einhver—frá forsendum aðfallsgreiningar sýnir þessi mynd?


ANOVA

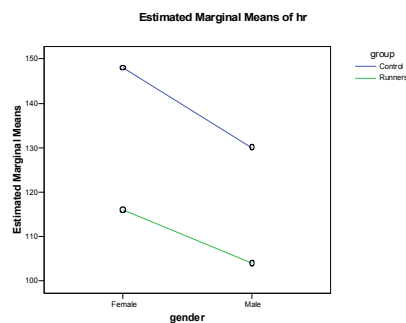
Model	Sum of Squares	ddf	Mean Square	F	Sig.
1 Regression	1302,08	1
Residual	961,00	36		
Total	37		

6. Í þessa dreifgreiningartöflu vantar mikið af upplýsingum. Settu réttar upplýsingar í alla reitina sem hafa punktalínur í stað talna.
7. Gerðu ráð fyrir að þú viljir athuga tengsl hjúskaparstöðu (einhleypur, í sambúð) og starfsaldurs (í árum) við heildarlaun starfsmanna stórs sveitarfélags. Útskýrðu með tilbúnum dæmum sem tilgreina launamun eftir hjúskaparstöðu, starfsaldri eða eftir báðum breytum samtímis hvernig *línuleg tengsl* þeirra við heildarlaun ættu að birtast.
8. Í skilaverkefni 2 var hallastuðullinn fyrir heilastærð (MRIPix) jafnt og 0,00012 þegar hún var eina frumbreytan en hækkaði upp í 0,00021 þegar líkamshæð var einnig inni í líkaninu. Hvað er á seyði: (a) Hvers vegna er hallastuðullinn tvöfalt hærri í síðara tilvikinu og (b) nákvæmlega hvaða merkingu hefur hallastuðullinn í hvoru tilviki fyrir sig?
9. Hvaða áhrif hefur það réttmætri niðurstaðna aðfallsgreiningar ef frumbreyturnar eru allar með miklum gólfhrifum (*ceiling effect*), dreifing þeirra mikið skekkt og með mörg frávíksgildi (frávillinga)?


10. Hversu hátt þarf R^2 að til að eitthvað sé að marka niðurstöðu aðfallsgreiningar?
11. Hvaða núlltilgátu athugar Próf Levenes, nákvæmlega hvaða forsendu dreifigreiningar tengist hún og hvaða aðrar aðferðir er mælt með að nota til að kanna þá forsendu?
12. Á hvern hátt gefur F -próf dreifigreiningar takmarkaðar upplýsingar um meðaltöl hópanna í þýði? Hvernig má bregðast við því?
13. Hver er meginmunurinn á fyrirfram (*a priori*) og eftir á (*post hoc; multiple comparisons*) samanburðum? Hvaða áhrif hefur þessi munur á tíðni höfnunarmistaka (*Type I error rate*)?
14. Í hverju er aðferð Bonferronis fólgin og hvaða vanda er henni ætlað að leysa?
15. Hvað nákvæmlega felst í samvirkni (*interaction*) tveggja frumbreyta? Komdu einnig með skýrt dæmi um samvirkni þar sem þú skýrir nákvæmlega hvers vegna þetta er samvirkni en ekki meginhrif (*main effect*).
16. Hvað nákvæmlega er fylgnileif (*residual correlation*) í þáttagreiningu og að hvaða leyti og hvernig er hægt að styðjast við hana við úrvinnsluna?



17. Hvaða frávik—ef einhver—frá forsendum dreifigreiningar sýnir þessi mynd?
18. Sæktu skrána eg13_011.por í heimaprófsmöppuna í Uglu. Þetta eru gögnin sem voru notuð í dæmi 13.8 í bókinni. Lestu vel lýsinguna á gögnunum í bókinni (bls. 813). Lestu gögnin inn í SPSS með því að nota File/ Open / Data.... Þar undir Files of type velur þú SPSS Portable (`.por`). 



- i. Endurgerðu myndina sem er hér fyrir ofan.
 - ii. Túlkaðu niðurstöðurnar sem sjást á myndinni í samræmi við lýsinguna á gagnasafninu í bókinni.
19. Hvaða upplýsingar gefur R^2 í aðfallsgreiningu?


20. Nefndu þrjár mikilvægar forsendur aðfallsgreiningar og gerðu grein fyrir því hvernig sé best að kanna hvort viðkomandi forsendur séu fyrir hendi?
21. Hvað nákvæmlega er misleitni (*heteroscedasticity*) og hvaða aðferðir henta best til að meta hana í tvíhliða dreifigreiningu?
22. Hvaða hlutverki gegnir snúningur (*rotation*) í þáttgreiningu? Hver er munurinn á hornréttum (*orthogonal*) og hornskökku (*oblique; nonorthogonal*) snúningi.
23. Sæktu skrána CSDData.por í heimaþrófsmöppuna í Uglu. Lestu vel lýsinguna á gögnunum í bókinni (bls. D2). Lestu gögnin inn í SPSS með því að nota File/ Open / Data.... Þar undir Files of type velur þú SPSS Portable (*.por). 

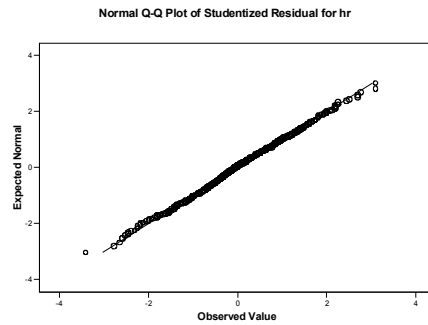
ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	26,303	2	13,151	26,626	,000 ^a
	Residual	109,160	221	,494		
	Total	135,463	223			

a. Predictors: (Constant), satm , hsm

b. Dependent Variable: gpa

- i. Framkvæmdu aðfallsgreiningu þar sem GPA er fylgibreytan og SATM og HSM eru frumbreytur. Þú átt að fá m.a. töfluna hér fyrir ofan. Berðu þínar niðurstöður saman við hana og gakkðu úr skugga um að þú hafir framkvæmt úrvinnsluna rétt.
 - ii. Gefðu upp aðfallsjöfnuna eins og hún er metin af forritinu.
 - iii. Túlkaðu alla hallastuðla. (a) Segðu hvað þeir merkja nákvæmlega svona tölfræðilega séð. (b) Útskýrðu með mannamáli hvað niðurstöðurnar þýða.
 - iv. Reiknaðu 95% öryggisbil fyrir hallastuðlana. (a) Hvaða ályktanir getur þú dregið um áhrif HSM? (b) En um SATM?
 - v. Segðu á einföldu mannamáli hvað niðurstöðurnar þýða? Miðaðu við að amma þín myndi skilja þig.
24. Á glærinni **Túlkun hallatalna í Marghliða aðfallsgreiningu** stendur eftirfarandi: „Áhrif einkunna í vísindum og ensku eru ekki marktæk. Þau eru einnig lítil, jafnvel þótt horft sé til efri marka öryggisbilanna.“ Hvað í ósköpunum er átt við?
- i. Halltölurnar eru ómarktækar og því skyldi ég þá skoða öryggisbilið?
 - ii. Hvaða máli skiptir það að efri mörk öryggisbilanna séu lág?
25. Á glærinni **Mynstur meðaltala í Einhliða dreifigreiningu** stendur eftirfarandi: „Allsherjarprófið í dreifigreiningu segir ekkert um mynstur meðaltala, þ.e. hver þeirra séu ólík og hver eins.“ Þetta er nú allt frekar dularfullt.
- i. Hvaða allsherjarpróf er þetta?
 - ii. Hvers vegna segir það ekkert til um mynstur meðaltala?
 - iii. Hvað getum við gert ef við viljum vita hvert mynstrið er?
26. Sæktu skrána eg13_011.por í heimaþrófsmöppuna í Uglu. Þetta eru gögnin sem voru notuð í dæmi 13.11 í bókinni. Lestu vel lýsinguna á gögnunum í bókinni (bls. 784). Lestu gögnin inn í SPSS með því að nota File/ Open / Data.... Þar undir Files of type velur þú SPSS Portable (*.por). 



- i. Framkvæmdu tvíhliða dreifigreiningu og endurgerðu myndina hér fyrir ofan.
- ii. Túlkaðu myndina með sérstakri áherslu á það hvort forsendur úrvinnslunnar séu til staðar.

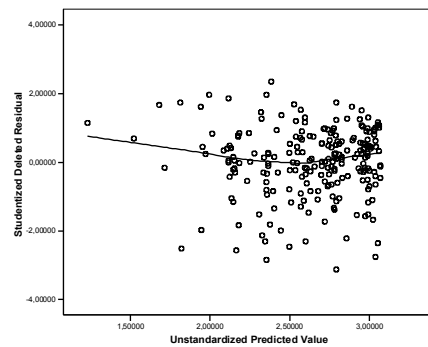
27. Sæktu skrána CSData.por í heimaprófsmöppuna í Uglu. Lestu vel lýsinguna á gögnunum í bókinni (bls. D3). Lestu gögnin inn í SPSS með því að nota File/ Open / Data.... Þar undir Files of type velur þú SPSS Portable (*.por). **H**

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	26,303	2	13,151	26,626	,000 ^b
	Residual	109,160	221	,494		
	Total	135,463	223			

a. Predictors: (Constant), satm , hsm
 b. Dependent Variable: gpa

- i. Framkvæmdu aðfallsgreiningu þar sem GPA er fylgibreytan og SATM og HSM eru frumbreytur. Þú átt að fá m.a. töfluna hér fyrir ofan. Berðu þínar niðurstöður saman við hana og gakktu úr skugga um að þú hafir framkvæmt úrvinnsluna rétt.



- ii. Gerðu mynd eins og þessa hérna. Línan er búin til með því að biðja um Loess-línu með 80% glugga.

28. Í fyrirlestrinum **Marghliða aðfallsgreining** er á glærum 1–4 og 14–15 rætt um sérhæf og ósérhæf áhrif og skilyrta og óskilyrta hallastuðla. Það er einnig rætt stuttlega um þetta efni á bls. 695 í 5. útgáfu kennslubókarinnar. **H**

- i. Útskýrðu með *eigin orðalagi* hvers vegna óskilyrta hallastuðull HSS á glæru 14 er 0,15 meðan skilyrta hallastuðullinn er 0,03.
- ii. Útskýrðu með eins skýru orðalagi og þú getur hvað felst í því að áhrif sjálfsmýndar sé *leiðrétt* fyrir greind í líkani 2 á glæru 15. Gættu þess að nota þitt eigið orðalag.

29. Í fyrirlestrinum **Einhliða dreifigreiningu** er fjallað um misleitni (*heteroscedasticity*) á glærunni **Misleitni** og gerð grein fyrir forsendum dreifigreiningar á glærunni **Einhliða dreifigreining**. **H**

- i. Ef alvarleg misleitni væri til staðar, hvernig myndir þú búast við að myndin á **Misleitni** myndi líta út? Hvers vegna?
 - ii. Nákvæmlega hvaða forsendu dreifigreiningar skv. **Einhliða dreifigreiningu** er brotin ef misleitni er til staðar? Tilgreindu nákvæmlega *hvað* á glærunni gefur til kynna forsenduna um einsleitni (*homoscedasticity*).
30. Lestu vel lýsingu á rannsókninni á öryggi á vinnustað (*study of workplace safety*) sem lýst er í sýnidæmum (*example*) 12.3–12.21. Útskýrðu síðan eftirfarandi: **H**
- i. Nákvæmlega hvaða spurningar prófa samanburðirnir (*contrasts*) tveir sem lýst er í sýnidæmi 12.17? Svarið þarf að vera á einföldu og skýru *mannamáli* (*ordinary everyday language, please*)!
 - ii. Hvaða ályktun get ég dregið um ψ_2 á grundvelli tölfræðiprófsins í sýnidæmi 12.20 og öryggisbilsins í sýnidæmi 12.21? Notaðu þitt eigið orðalag, *ekki* éta upp úr bókinni!
 - iii. Hvaða kosti hefur það að nota slíka samanburði fremur en að byggja eingöngu á allsherjarprófinu (*omnibus test*) sem lýst er í upphafi sýnidæmis 12.14?
31. Í hvaða tilvikum notum við aðferð Bonferronis og nákvæmlega hvaða vanda er henni ætlað að leysa?
32. Í hverju felst skriðupróf Cattells (*Cattell's scree test*), hvernig er það notað og hvaða afleiðingar hefur það ef dregnir eru annað hvort færri eða fleiri þættir (*factors*) heldur en skriðuprófið gefur til kynna?
33. Hverjar eru helstu forsendur (*assumptions*) dreifigreiningar og hvernig er best að kanna hvort þær séu fyrir hendi?
34. Í hverju nákvæmlega felst aðferð Scheffé við eftir á samanburði? Fyrir hvers konar samanburði hentar aðferðin?
35. Í rannsókn á 282 konum var athugað hvaða þættir tengdust ánægju við líkamsrækt. Fylgibreytan, *Ánægja*, gat tekið gildin 18 til 136. Frumbreyturnar voru þrjár: Smekkur fyrir tónlistinni (*Tónlistarsmekkur*) var á kvarðanum 4 til 28; Ánægja með leiðbeinandann var á bilinu 6 til 42; *Samsemd*, þ.e. hvort viðkomandi fann sig í líkamsræktinni.
- Fylgni á milli breytanna fjögurra var umtalsverð og í öllum tilvikum marktæk með $p < 0,05$. Hallatalan fyrir Tónlistarsmekk var 1,02, hún var 0,96 fyrir Ánægju með leiðbeinandann og 0,30 fyrir Samsemd.
- Í niðurstöðum rannsóknarinnar kom eftirfarandi einnig fram: $R^2 = 0,33$.
- i. Geturðu gefið upp aðfallsjöfnuna? Ef svarið er jákvætt, skaltu gefa jöfnuna sem svar. Ef svarið er neikvætt, þarftu að útskýra hvers vegna.
 - ii. Allar fylgnitölur voru margtækar. Þýðir það að hallastuðlarnir séu marktækir líka? Rökstyddu svarið nákvæmlega.
 - iii. Núlltilgátan var að allar hallatölur væru núll. Niðurstaða tölfræðiprófsins var $F = 45,64$. Gefðu upp frígráður prófsins, marktækt þess og túlkaðu niðurstöðuna.
 - iv. Hve hátt hlutfall af dreifingu fylgibreytunnar má skýra með frumbreytunum þremur?
36. Sæktu skrána `spurning2.sav` í heimaprófsmöppuna í Uglu. Þessi skrá innheldur sömu gögn—`csdata.sav`—og eru notuð í fyrirlestrinum **Marghliða aðfallsgreiningu** nema hvað fylgibreytan er ný og heitir núna `newy`. **H**
- i. Framkvæmdu aðfallsgreiningu þar sem `newy` er fylgibreytan og HSS, HSE og HSM eru frumbreytur. Birtu hallastuðlana og dreifigreiningartöfluna.

- ii. Endurgerðu helstu myndritin í fyrirlestrinum **Marghliða aðfallsgreiningu** með þessum nýju gögnum: Kassarit af leif, normalrit af leif og leifarritið eins og á glæru 20.
 - iii. Tilgreindu fyrir hverja mynd hvaða eiginleikar villunnar hægt er að meta með þeirri mynd og lýstu því hvort gögnin standist forsendur dreifigreiningar eins og þeim er lýst t.d. á **Marghliða aðfallsgreiningu** og á bls. 685 í 5. útgáfu kennslubókarinnar (ISP) undir fyrirsögninni *Multiple linear regression model*.
37. Skoðaðu vel mynd 12.6 á bls. 728 í bókinni (einnig birt á glæru 1 í **Einhliða dreifigreiningu**) með hliðsjón af forsendunni um einsleitni (*homoscedasticity*). **H**
- i. Útskýrðu nákvæmlega (a) hvað átt er við með einsleitni, sbr. umræðu á bls. 728–729 og (b) tilgreindu nákvæmlega hvernig einsleitnin birtist á mynd 12.6.
 - ii. Lýstu þremur meginaðferðum við að meta einsleitni og útskýrðu hvernig nákvæmlega misleitni (*heteroscedasticity*) sést með aðferðunum hverjum fyrir sig.
38. Skoðaðu dæmi 12.18 á bls. 740 í bókinni. Þú gætir þurft að lesa blaðsíðurnar og dæmin á undan og eftir einnig. **H**

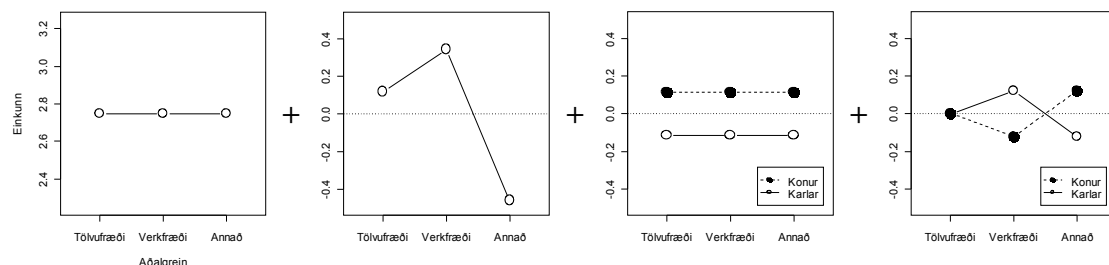
Gættu þín á villu í bókinni. Samanburðurinn á að vera svona:

$$c_1 = \bar{X}_{SU} - \frac{1}{2}(\bar{X}_{UN} + \bar{X}_{SK})$$

Þeir sem gerðu rannsóknina tóku eftir því að verkstjórar voru með hæstu niðurstöðuna, öruggasta vinnuumhverfið. Þeir ákváðu því að *endurtaka* rannsóknina og gera fyrirfram samanburð þar sem verkstjórar (*supervisors*) væru bornir saman við hina. Þetta rökstuddu þeir með því að verkstjórar væru með allt annað vinnuumhverfi en faglærðir og ófaglærðir, bæði fylgdust þeir með og stjórnudu í stað þess að vinna verkin sjálfur auk þess sem umtalsverður hluti vinnutímans eru fundir og skrifborðsvinna.

Niðurstaða samanburðarins var tæpum 7,9 stig verkstjórum í hag. Þeir birtu niðurstöðu sína svona: „Verkstjórar mátu vinnuumhverfi sitt öruggara ($M= 78,8$) heldur en aðrir starfsmenn ($M= 71,0$), $t(587)= 3,04$. $p < 0,01$.“ Þeir birtu ekki allsherjarprófið (*omnibus test*).

- i. Útskýrðu hvers vegna þeir máttu nota fyrir fram (*a priori; contrasts*) prófi í þessari nýju rannsókn. Gefðu skýrt og nákvæmt svar.
- ii. Hverjir eru kostir fyrir fram prófs og hverjir eru ókostir eftir á (*post hoc; multiple comparisons*) prófs í þessu samhengi?
- ii. Hvernig stendur á því að þeir birta ekki allsherjarprófið, er það réttmætt?



39. Skoðaðu vel ofangreinda mynd sem er sýnidæminu úr fyrirlestrinum **Tvíhliða dreifigreiningu** og sýnir sundurliðun meðaltala af glærunni **Mat á samvirkni**, sbr. neðri töfluna á glærunni.
- i. Útskýrðu hverja mynd fyrir sig, þ.e. nákvæmlega hvað hún sýnir.

- ii. Berðu aðra mynd frá vinstri saman við myndina á glærunni Samvirkni. Í báðum tilvikum er hægt að sjá meginhrif. Að hvaða leyti er myndin hér skýrari heldur en myndin á glærunni?
 - iii. Berðu síðustu myndina saman við myndina á glærunni Samvirkni. Í báðum tilvikum sést samvirkni. Að hvaða leyti er myndin hér skýrari heldur en myndin á glærunni eða er það svo að þessi mynd sé skýrari?
40. Hvað er áhættuhlutfall (*odds ratio*) og nákvæmlega hvaða upplýsingar gefur það okkur um tengsl frum- og fylgibreytu?
41. Marghliða aðfallsgreining (*multiple regression*) gefur sérhæf áhrif frumbreytu á fylgibreytu þar sem leiðrétt er fyrir (*controlled for*) áhrif annarra frumbreyta sem eru í líkaninu.

Hvað nákvæmlega felst í þessum sérhæfu eða leiðréttu áhrifum? Hvernig fæ ég óleiðrétt (ósérhæf) áhrif og hvernig yrðu þau öðru vísi, þ.e. stærri eða minni, og hve miklu munar? Skýrðu nákvæmlega m.a. með því að taka ímyndað dæmi.

42. Athuguð voru tengsl einkunna þar sem einkunn í Tölfræði II var fylgibreytan en einkunn í Tölfræði I frumbreyta. Í ljós kom að hallastuðullinn var 0,57. Þegar meðaleinkunn á stúdentsprófi var bætt við sem frumbreytu lækkaði hallastuðull fyrir Tölfræði I verulega eða niður í 0,36. [Þetta eru skáldaðar tölur!]
- i. Túlkaðu fyrri hallastuðulinn, þ.e. hallastuðul fyrir Tölfræði I þegar sú einkunn er eina frumbreyta líkansins.
 - ii. Túlkaðu seinni hallastuðulinn, þ.e. hallastuðul fyrir Tölfræði I þegar bæði sú einkunn og stúdentsprófseinkunn eru frumbreytur í líkaninu.
 - iii. Útskýrðu (a) hvers vegna seinni hallastuðullinn er lægri og (b) hvers vegna meðaleinkunn á stúdentsprófi breytir hallastuðli einkunnar í Tölfræði I.
43. Sæktu skrána `spurning2.sav` í heimaprófsmöppuna í Uglu. Þessi skrá á uppruna sinn í töflu 1.9 í kennslubókinni en gögnum hefur verið lítillega breytt. **H**
- i. Framkvæmdu aðfallsgreiningu þar sem `gpa` er fylgibreytan og `iq`, `gender` og `concept` eru frumbreytur. Birtu hallastuðlana og dreifigreiningartöfluna.
 - ii. Taktu afstöðu til þess hvort forsendur aðfallsgreiningar standist með því að skoða leifina. Birtu kassarit af leif, normalrit af leif og leifarrit og túlkaðu þessi rit með hliðsjón af forsendunum.
 - iii. Reiknaðu spágildi fyrir einkunn (`gpa`) miðað við ólík gildi sjálfsmýndar (`concept`). Gerðu þetta sérstaklega fyrir karla (hafa gildið 2 á `gender`) og fyrir konur (hafa gildið 1) og miðaðu við að í báðum tilvikum sé greindarvísitalan (`iq`) nákvæmlega 100,0. Birtu niðurstöðurnar myndrænt, þ.e. ein mynd þar sem sjálfsmýnd er á lárétta ásnum (20–80), einkunn á lóðrétta ásnum (4,0–7,0) og spágildi fyrir konur eru sýnd með rauðri línu en spágildi fyrir karla með blárrí (eða með sambærilegri aðgreiningu eftir kyni).

Levene's Test of Equality of Error Variances			
Dependent variable: Tekjur síðasta sumar			
F	df1	df2	Sig.
4,803	3	45	0,005

44. Skoðaðu vel töfluna með prófi Levenes hér fyrir ofan og svaraðu eftirfarandi spurningum.
- i. Hvað er verið að prófa með þessu prófi og hver er núlltilgáta þess.

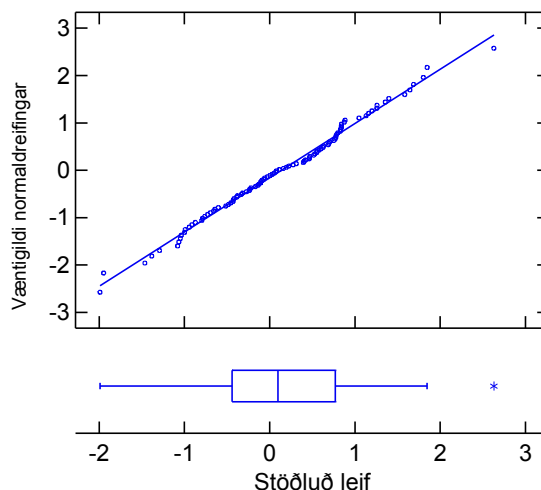
- ii. Hvaða ályktun megum við draga af annars vegar marktækri og hins vegar ómarktækri niðurstöðu?
 - iii. Hvaða aðrar aðferðir höfum við til að meta það sem verið er að prófa og hvernig samþæfum við niðurstöður þeirra við þessar?
45. Á bls. 739 í kennslubókinni segir: „A contrast expresses an effect in the population as a combination of population means. To estimate the contrast, form the corresponding **sample contrasts** by using sample means in place of population means. Under the ANOVA assumptions, a sample contrast is a linear combination of independent normal variables and therefore has a normal distribution.“
- i. Útskýrðu ofangreint með þínum eigin orðum. Gættu þess að skýrt komi fram (a) hvað felist í fyrirfram samanburði, (b) hver er munurinn á samanburði í úrtaki og samanburði í þýði, (c) í hvaða skilningi samanburðir eru línulegir og (d) hvernig *samanburðir* geta haft normaldreifingu.
45. Brussella gerði dreifigreiningu á meðaleinkunnum í samræmdum prófum í 170 íslenskum grunnskólum. Þar sem hún vildi vita hverjir þeirra væru ólíkir, bar hún þá saman tvo og tvo og reiknaði *t*-próf milli meðaltalanna. Hún fann 20 þör skóla sem voru með marktækt ólíkar meðaleinkunnir og í sumum tilvikum var munur mjög mikill.

Hægfinnur var ekki hrifinn af niðurstöðunum og hafnaði þeim alfarið. Hann sagði að Brussella hafði ranglega notað LSD-próf (sjá bls. 744) og niðurstaðan gæti verið fyrir tilviljun eina. Nota þyrfti leiðréttingu eins og t.d. Bonferroni eða Scheffé.

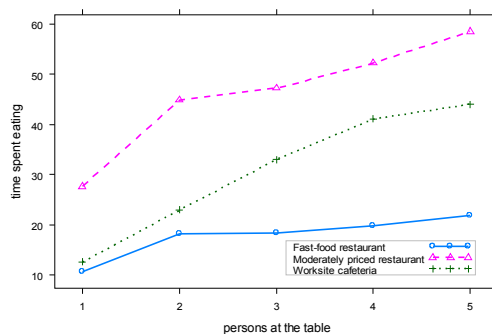
Brusella snöggreiddist: „Þetta eru sko engar ofskeyningar hjá mér og það væri fráleitt að fara að leiðrétta niðurstöðuna því þá myndi marktæktin hverfa. Sjálfur ert þú bara aumkunarverður hasshaus.“

Útskýrðu hvað Hægfinnur á við með sínum athugasemdum.

- i. Hvernig getur þetta verið fyrir tilviljun þrátt fyrir marktæktina?
 - ii. Hvað kemur fíkniefnið LSD málinu við eða hvað er maðurinn að tala um í sambandi við LSD?
 - iii. Hver er munurinn á villutíðninni α (alfa) sem *t*-próf miðast við og þeirri villutíðni sem í reynd er stuðst við ef leiðrétt er með Bonferroni eða Scheffé (sjá bls. 742–748)?
46. Í hverju felast eftir á samanburðir (*post hoc comparisons; multiple comparisons*) og hvaða vandi fylgir því að gera marga eftir á samanburði þar sem hver og einn miðast við α ?



47. Myndin sýnir normalrit (*probability plot*) og kassarit (*box plot*) fyrir staðlaða leif (*standardized residual*) aðfallsgreiningar (*regression*). Segðu mér tvennt: (a) Hvaða frávik (*deviation*) frá forsendum (*assumptions*) aðfallsgreiningar sjást almennt (*can in general be detected*) á normal- og kassaritum? (b) Sjást einhver þessara frávik (*can any such deviations be seen*) á myndinni? Rökstyddu og vertu nákvæm og skýr í framsetningu.
48. Sæktu skrána `happiness.sav` í heimaprófsmöppuna í Uglu. Þessi skrá er einnig að finna á geisladiskinum sem fylgir kennslubókinni og er notuð í dæmum 11.31–11.33. Lestu vel lýsinguna á gagnasafninu í viðauka D4 í kennslubókinni og kynntu þér upplýsingar sem koma fram í dæmunum sem nefnd eru hér að ofan. **H**
- Gerðu marghliða aðfallsgreiningu þar LSI er fylgibreytan en DEMOCRAC, GINI og LIFE eru frumbreytur. Túlkaðu niðurstöðuna með tilliti til hallastuðlanna.
 - Gerðu aðra aðfallsgreiningu þar sem breytan CORRUPT bætist við sem frumbreyta. Túlkaðu hallastuðlana hérna einnig.
 - Reiknaðu út hver meðallífshamingja þjóðar ætti að vera samkvæmt líkönunum tveimur miðað við að DEMOCRAC taki gildin 0, 1, 2, 3, 4, 5 og 6 en aðrar breytur séu settar jafnar meðaltölum sínum. Birtu þessar niðurstöður sem mynd þar sem stig lýðræðis (DEMOCRAC) er á lárétta ásnum en spáð lífshamingja (LSI) er á lóðrétta ásnum. Myndin ætti að hafa eina línu fyrir fyrra líkanið og aðra fyrir seinna líkanið.
 - Útskýrðu (a) hvernig línurnar eru ólíkar og (b) hvers vegna hallastuðullinn fyrir DEMOCRAC breytist milli líkana. Athugaðu að það er ætlast til þess að þau skiljir í hvaða grundvallaratriðum líkönin eru ólík og sért fær um að túlka þann mun á skýran og skiljanlegan hátt.
49. Útskýrðu hvaða upplýsingar fást með R^2 í aðfallsgreiningu. Miðaðu við að svara eftirfarandi spurningum skýrt og afdráttarlaust.
- Hvaða er þessi stuðull að mæla nákvæmlega?
 - Hvað þarf R^2 að vera hátt til þess að það sé að marka niðurstöðu aðfallsgreiningar? Gættu þess að ræða þetta og rökstyðja ítarlega!
50. Á einni glæru námskeiðsins stendur eftirfarandi: „Dreifigreining miðar við að villan hafi sama staðalfrávik í öllum hópum.“
- Þýðir þetta að það verði að vera sama staðalfrávik í öllum hópum rannsóknarinnar í úrtakinu? Útskýrðu og rökstyddu vandlega!
 - Teiknaðu (handteiknuð mynd er fullnægjandi) kassarit þar sem þessi forsenda (*assumption*) er augljóslega brotin.
 - Hvaða aðferðir höfum við til að meta þessa forsendu og hvernig samhæfum við niðurstöður þeirra?



51. [Bell og Pliner \(2003\)](#) gerðu athugun á því hvernig fjöldi þeirra sem situr við borð hefur áhrif á hversu lengi er setið við máltíð.

Þeir fundu meginhrif fyrir fjölda þeirra sem matast var með, $F(4, 352) = 58,0$, $p < 0,001$, þannig að matartíminn varð stystur (17 mínútur) þegar aðeins einn sat til borðs, tæplega 33 mínútur þegar þrír sátu til borðs og rúm 41 mínúta þegar fimm (eða fleiri) sátu til borðs. Þeir fengu einnig meginhrif fyrir þann stað sem matast var á, $F(2, 352) = 232,6$, $p < 0,001$, þar sem matartíminn varð lengstur (46 mínútur) þegar matast var á miðlungsdýrum veitingarstað, næstlengstur (31 mín.) í mötuneyti og stystur (18 mín.) þegar matast var á skyndibitastað.

Það flækti málið eilítið að í ljós kom samvirkni ofangreindra tveggja breyta, $F(8, 352) = 7,7$, $p < 0,001$. Myndin sýnir samvirkni breytanna.

- i. Gerðu nákvæma grein fyrir vandanum við að meta meginhrif (*main effect*) þegar samvirknihrif (*interaction effect*) eru til staðar.
- ii. Í hverju eru samvirknihrifin nákvæmlega fólgin? Gættu þess að í skýringu þinni komi skýrt fram bæði almenn útskýring á samvirknihrifum og útlitun á samvirknihrifunum eins og þau birtast á myndinni hér að ofan.
- iii. Er í þessu tiltekna tilviki merkingarbært að túlka meginhrifin? Útskýrðu nákvæmlega og gættu þess sérstaklega að túlka hvor meginhrif fyrir sig ef þú telur þau túlkanleg eða tilgreina að öðrum kosti nákvæmlega í hvoru tilviki fyrir sig í hvaða skilningi þau eru ótúlkanleg.
- iv. Ef þú telur meginhrifin vera túlkanleg, rissaðu þá upp samsvarandi mynd þar sem meginhrifin væru ótúlkanleg. Ef þú hins vegar telur meginhrifin vera ótúlkanleg, rissaðu þá upp samsvarandi mynd sem sýnir samasvarandi túlkanleg meginhrif.