

Arbetslogg 2016 vecka 11

Erik Bulow

14 mars 2016

Contents

1	Förberedelser	1
2	2016-03-14	2
2.1	Läsning av (Olkin and Finn 1995)	2
2.2	Läsning av (Yin and Fan 2001)	2
2.3	Läsning av (Bobko 2001)	5
2.4	Läsning av (Shieh 2010)	6
2.5	Läsning av (Wang and Thompson 2007)	6
3	2016-03-15	7
4	2016-03-17	8
5	2016-03-18	8
5.1	Läsning av (Croux and Dehon 2003)	8
5.2	Läsning av (Renaud and Victoria-Feser 2010)	8
5.3	Läsning av (Asuero, Sayago, and González 2006)	9
5.4	Läsning av (Hössjer 2008)	9
5.5	Läsning av (Burton et al. 2006)	9
	Referenser	10

1 Förberedelser

```
# Try it out!  
memory.limit(50000)
```

```
## [1] 50000
```

```
options(samplemetric.log = TRUE)  
set.seed(123)
```

2 2016-03-14

2.1 Läsning av (Olkin and Finn 1995)

Utvecklar metoder för att skapa konfidensintervall runt R^2 .

Vid första ögonkastet tycks artikeln fokuserad på ρ men så är inte fallet. Ställer upp ett par olika situationer där man vill jämföra olika typer av ρ^2 , t ex multiple mot simple och i olika samples etc. Teoretisk ansats men med tillämpning på riktig data.

OBS! Bygger på normalfördelningsantagande som kanske funkar för stora stickprov etc men det är vi ju egentligen inte intresserade av! Skriver i conclusion att deras metoder kan rekommenderas med viss försiktighet för $n \in [60, 200]$ och först därefter med större säkerhet.

2.2 Läsning av (Yin and Fan 2001)

Jämför olika formler för adjusted R^2 (nämner även empiriska metoder men fokuserar inte på det). Slutsats att (Wherry 1931) inte är den bästa metoden utan de rekommenderar istället Pratt eller Browne. Skriver dock att (Wherry 1931) är den mest använda formeln i SAS och SPSS (och vi vet ju själva att den också används i R). De undersöker empiriskt med monte-carlo.-simulering och tycks ha ett väldigt brett spektrum av paramtervärden att slumpa utifrån.

Skriver att olika studier visar olika resultat för vilken justering som är den bästa. Finns t ex resultat som pekar på att Browne bäst för just multipel korrelation.

Poängterar att man i mpnga studier blandat ihop (Ezekiel and Mordecai 1929) och (Wherry 1931) och iobland även ρ^2 med ρ_c^2 , vilket leder till felaktiga resultat. Nämner också att olika studier jmftr olika metoder, vilket gör att man saknar helhetsbild samt att utvärdering mha olika data set inte nödvändigtvis heller underlättar jämförelser. Förespråkar att använda genererad syntetisk data.

Nämner att R^2 används både för förståelse mn också för prediktion men att den överskattas i båda fallen, dvs att man tenderar få et lägre värde då samma modell tillämpas på nytt data set.

Skriver att man i princip alltid utgår från fixed modell i regression även om det finns teoeri för random models. Denna teori är dock alltför komplex för att ha fått ngt praktiskt genomslag. Detta bidrar dock också till överskattning av ρ^2 då det innebär att man ignorerar en källa till variation. Och denna variation ökar desutom extra mycket med fler variabler som antas fixa men som eg inte är det.

Nämner att om R_c, ρ_c baseras på samma modell men på dett nytt data set gäller: $E[R_c] \approx \rho_c < \rho < E[R]$

Samplpe multilpe correlation coefficient används som regel både för R och R_c . Det är dock känt att överskattningen ökar och att vi därmed måste man "shrink" eller "correct" R för att åstadkomma adjusted R .

Nämner att det finns åtmin 15 olika justeringsformler.

Menar här liksom även tidigare källa att Smiths formel enl (Ezekiel and Mordecai 1929) och beskrivet som så även av (Wherry 1931):

$$\hat{R}^2 = 1 - \frac{N}{N-p}(1 - R^2)$$

Även den formell som ibland kallas Wherrys formell 1:

$$\hat{R}^2 = 1 - \frac{N-1}{N-p-1}(1 - R^2)$$

föreslås av Ezekei.

Om dessa formler skrivs:

[...] cited with different names, listed here in decreasing order of frequency: the Wherry formula (Ayabe, 1985; Kennedy, 1988; Krus & Fuller, 1982; Schmitt, 1982; Stevens, 1996), the Ezekiel formula (Huberty & Mourad, 1980; Kromrey & Hines, 1996), the WherryMcNemer formula (Newman et al., 1979), and the CohenCohen formula (Kennedy, 1988). The Wherry formula-2 was also cited in one study as an estimator for cross-validation (Kennedy, 1988). This formula is currently being implemented by popular statistical packages for computing the adjusted R² in multiple regression procedures (e.g., SAS/STAT User's Guide, 1990; SPSS User's Guide, 1996).

Wherry formula-2:

$$\hat{R}^2 = 1 - \frac{N-1}{N-p}(1-R^2)$$

Denna formel presenterades verkligen av (Wherry 1931) men har i sin tur också kallats till exempel McNemer formula och den har också misstagits för Wherry-1.

Tre olika approximationer till (Olkin and Pratt 1958) redovisas och om denna skrivs:

These formulas were cited as the Olkin and Pratt formula in several studies (Ayabe, 1985; Claudy, 1978; Huberty & Mourad, 1980; Krus & Fuller, 1982) and were cited as the Herzberg formula in one study (Cummings, 1982).

Det som sedan kallas Pratt's formula är ytterligare en approximation av (Olkin and Pratt 1958) och som är väldigt lik de tidigare.

$$\hat{R}^2 = 1 - \frac{(N-3)(1-R^2)}{(N-p-1)} \left[1 + \frac{2(1-R^2)}{N-p-2.3} \right]$$

Denna formel presenterades första gången i personlig kommunikation men finns beskriven i bl a (Claudy 1978).

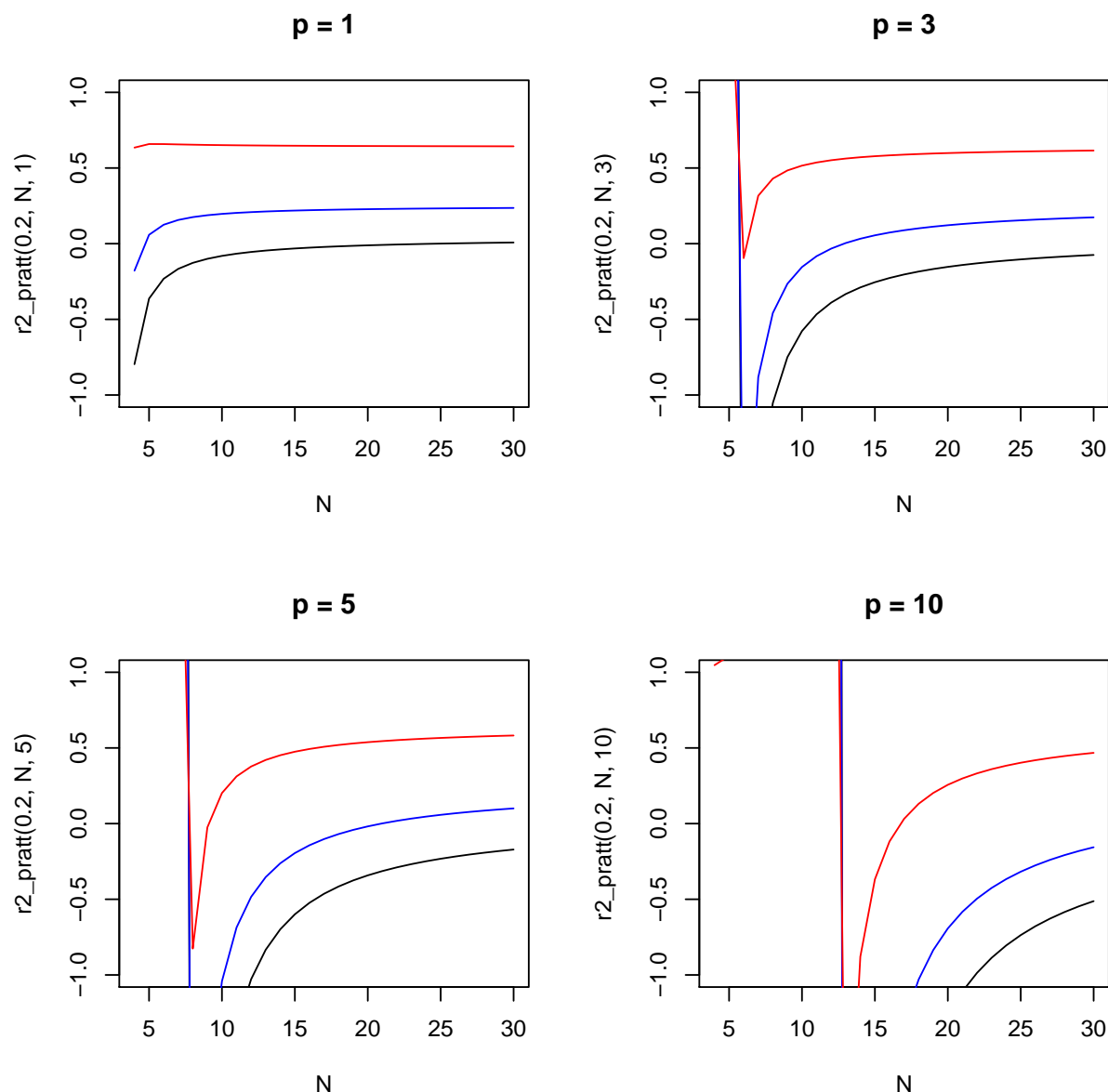
```
r2_pratt <- function(R, N, p) {
  1 - (
    ((N - 3) * (1 - R ^ 2) / (N - p - 1)) *
    (1 + (2 * (1 - R ^ 2)) / (N - p - 2.3))
  )
}
N <- 4:30

par(mfrow = c(2,2))
plot(N, r2_pratt(.2, N, 1), type = "l", ylim = c(-1, 1), main = "p = 1")
lines(N, r2_pratt(.5, N, 1), type = "l", col = "blue")
lines(N, r2_pratt(.8, N, 1), type = "l", col = "red")

plot(N, r2_pratt(.2, N, 3), type = "l", ylim = c(-1, 1), main = "p = 3")
lines(N, r2_pratt(.5, N, 3), type = "l", col = "blue")
lines(N, r2_pratt(.8, N, 3), type = "l", col = "red")

plot(N, r2_pratt(.2, N, 5), type = "l", ylim = c(-1, 1), main = "p = 5")
lines(N, r2_pratt(.5, N, 5), type = "l", col = "blue")
lines(N, r2_pratt(.8, N, 5), type = "l", col = "red")

plot(N, r2_pratt(.2, N, 10), type = "l", ylim = c(-1, 1), main = "p = 10")
lines(N, r2_pratt(.5, N, 10), type = "l", col = "blue")
lines(N, r2_pratt(.8, N, 10), type = "l", col = "red")
```



Vi ser av bilderna att vi också borde ha en lägre gräns för $/p$ kanske för att det alls ska vara meningsfullt att titta på detta.

Kallar (Claudy 1978):s formel för Claudy formula-3.

Identifierar också 9 st formler för R_c (cross-validity coefficient) som jag inte återger lika noggrant. En del av dessa formler utgår i sin tur från ρ^2 men skattar då denna i sin tur via atningen (Wherry 1931) eller (Olkin and Pratt 1958).

Här säkuleras även kring att en formel kan ha missförståts pga tryckfel (blandat ihop + och -).

Simulerar för $\rho = .2, .5, .8, p = 2, 4, 8, N = 20, 40, 60, 100, 200$.

Nämner att (Claudy 1978) konstaterat att multikolaritet inte haft ngn större effekt i sammanhanget men detta tas ändå med som en faktor i denna undersökning. Intercorrelation mellan independent variables sattes till .10, .30, .50 (samma för alls kombinationer av variabler).

Resulterade i designt experiment med $3 \times 3 \times 5 \times 3 = 135$ kombinationer. Upprepades 500 ggr per kombination => 67500 replicates totalt.

All data multivariat normalfördelat.

Beskriver samplingsproceduren i SAS.

Jämförde mean och sd för alla metoder baserat på de 500 repetitionerna. Antar att skattning unbiased om mean inom $R \in [\rho \pm 0.01]$. Presenterar andel unbiased för olika metoder. Uppställningen är tydlig och klar och här framkommer att Pratt bäst och att den vanliga formeln inte särskilt bra. Bäst resultat ges då N/p stor.

Observera att (Skidmore and Thompson 2011) refererar även till denna artikel men att deras resultat är att den vanliga formeln är OK men även de menar att Pratt egentligen är bäst! Skidmore har även $N = 10$ och inkluderar även skeva fördel Osäker på hur denna skillnad uppstår. En skillnad är förstås att Skidmore inte bara tittar på normalfördelat data utan även skeva fördelningar etc. Man gör även lite noggrannare analys av bias (mer än att bara konstatera huruvida ett estimat kan betraktas som unbiased eller inte).

Här gjordes också en variansanalys av bias med slutsatser. Både enskilda och interaktionssammanslagna faktorer visar sig kunna förklara väldigt små andelar av den totala variationen men viktigaste till minst viktiga är $N, \rho^2, N\rho^2, p, Np, N\rho^2p$. Andelen förklarad varians är större för de empiriska metoderna. Allra viktigast är dock förhållandet N/p och inte ngn enskild parameter.

Skriver att den vanliga formeln lika bra som övriga endast då $N/p \approx 100$. (Skidmore and Thompson 2011) är alltså lite snällare och säger att den vanliga formeln är OK (men säger inte att den skulle vara jättebra).

Observera att det enda vi kollar i denna artikel ju är andelen unbiased enl def medan Skidmore undersöker mer än så (vilket de också själva påpekar).

På det hela taget en välskriven och mkt intressant artikel.

2.3 Läsning av (Bobko 2001)

OBS! Om ρ , inte ρ^2

Detta är ett bokkapitell ur en bok som ev är ganska grundläggande och som handlar enbart om correlation och regression. Kanske kan vara en poäng att referera till denna för en allra första introduktion för ovana läsare?

Gör skillnad på två olika skrivsätt där:

$$r_{X,Y} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \sum (Y - \bar{Y})^2}}$$

kallas conceptual formula och där dess algebraiska motsvarighet:

$$r_{X,Y} = \frac{\sum XY - (\sum X)(\sum Y)/n}{\sqrt{[\sum X^2 - (\sum X)^2/n][\sum Y^2 - (\sum Y)^2/n]}}$$

kallas computational formula. Den första lättare att tolka, den andra att beräkna.

Påpekar att r inte är robust utan att outliers kan påverka dess värde oproportionerligt mkt. Detta är väl antagligen också värt att nämnas i sammanhanget då små stickprov diskuteras.

Jag läser inte färdigt denna text!

2.4 Läsning av (Shieh 2010)

Kommer här fram till att vanliga r trots allt kan vara bättre än unbiased versioner. Obs ρ , ej ρ^2 . Undersöker:

$$RMSE = \sqrt{MSE(\hat{\rho}, \rho)} = \sqrt{E[(\hat{\rho} - \rho)^2]} = Bias(\hat{\rho}, \rho)^2 + Var(\hat{\rho})$$

Fig 1 och 2 ger rätt snygga illustrationer över förhållande mellan ρ och bias resp RMSE. Lätt att se mönstret! Olika linjer för olika n , $n = 20, 50, 100$.

Tycker att sample r duger för $|\rho| \leq .6$. Detta motiveras bl a med att det är beräkningsmässigt fördelaktigt jmfrt med olika adjusted versioner.

2.4.1 För ρ^2

OBS! Gäller r^2 , ej R^2 , dvs simple correlation, ej multiple.

$Bias(r^2, \rho^2) > 0$ for $0 \leq \rho \leq .70$ and $Bias(r^2, \rho^2) < 0$ for $\rho \geq .75$.

Finner att för adjusted enl (Wherry 1931) $\hat{\rho}_E^2$ (egen beteckning) och enl Pratt $\hat{\rho}_{PA}^2$ gäller (då $p = 1$):

According to these findings, $\hat{\rho}_E^2$ is advantageous in MSE for small $\rho < .30$, r^2 dominates for $.30 \leq \rho \leq .85$ and $\hat{\rho}_{PA}^2$ performs best for large $\rho > .85$.

2.5 Läsning av (Wang and Thompson 2007)

Denna artikels upplägg liknar väldigt mkt (Skidmore and Thompson 2011) men för lite färre formler. Dock finns även här olika fördelningar med olika kurtosis etc. Skidmore refererade till denna artikel och utgav sig just för att vara en förbättring i förhållande till denna.

Slutsats att R^2 bara är marginellt bias och att Smiths och (Ezekiel and Mordecai 1929) duger bra för korrigering.

Poängterar att många tidsskrifter nu kräver rapportering av effect size. Skriver att alla effektmått baserade på OLS kommer att överestimera sitt sanna värde då man får en sorts over fitting ty man får in sample bias.

Så här kan man citera de olika justerade:

For example, in the R^2 arena, the six primary correction candidates were proposed by Ezekiel (1929, 1930), Smith (as cited in Ezekiel, 1929, p. 100), Wherry (1931), Olkin and Pratt (1958), Pratt (personal communication, October 20, 1964, cited in Claudy, 1978), and Claudy (1978).

Annan bra förklarande text till varför storleken av ρ påverkar skattningen av sig själv:

Although the reasons why sample size and the number of measured variables impact sampling error are intuitively straightforward, least obvious is the reason population effect size impacts sampling error in estimating effect sizes. Thompson (2002) explained, As an extreme heuristic example, pretend that one was conducting a bivariate r^2 study in a situation in which the population r^2 value was 1.0. In this population scattergram, every person's asterisk is exactly on a single regression line. In this instance, even if the researcher draws ridiculously small samples, such as $n = 2$ or $n = 3$, and no matter which participants are drawn, the researcher simply cannot incorrectly estimate the variance-accounted-for effect size. That is, any two or three or four people will always define a straight line in the sample scattergram, and thus r^2 will always be 1.0. (p. 68)

3 2016-03-15

Möte med SN:

1. Artikeln bör vara då fokuserad som möjligt på ett “budskap”
2. Används de referenser som känns relevanta men inkludera inte ref bara för sakens skull.
3. Gör simulering med bivariat normalfördelad data och $\rho^2 = .1, .25, .5, .9, N = 5, 10, 20, 30, 40, 50, 100, 150, 200$ upprepa 1000 ggr. 4 grafer (2×2) med stickprov på x-axeln och ρ^2, R^2 på y-axeln. Olika linjer för de sex sedan tidigare presenterade justerade R^2 -värdena, ojusterade samt väntevärdet enl

Implementerar funktioner för medelvärde och varians av R^2 enl (Wishart, Kondo, and Elderton 1931):

```
# mean av  $R^2$  enl (i)
mean_r2 <- function(rho2, n, p) {
  1 - ((n - p) / (n - 1)) * (1 - rho2) * Re(hypergeo::hypergeo(1, 2, .5 * (n + 1), rho2))
}

# var av  $R^2$  enl (ii)
var_r2 <- function(rho2, n, p) {
  (
    (((n - p) * (n - p + 2)) / ((n - 1) * (n + 1))) *
    (1 - rho2) *
    Re(hypergeo::hypergeo(2, 2, .5 * (n + 3), rho2))
  ) - (1 - mean_r2(rho2, n, p)) ^ 2
}
```

Implementerar även funktioner för olika justerade R^2 -värden: smith, ezekiel, wherry, olkin_pratt1, pratt, claudy enl (Yin and Fan 2001). olkin_pratt enl (Olkin and Pratt 1958). olkin_pratt2 enl (Cattin 1980) där olkin_pratt1 och 2 är approximationer av olkin_pratt

```
adjusted_r2 <- function(r2, n, p, adj = c("smith", "ezekiel", "wherry", "olkin_pratt", "olkin_pratt1",
                                           "olkin_pratt2", "pratt", "claudy")) {

  smith      <- function() 1 - ((n - 1) / (n - p)) * (1 - r2)
  ezekiel    <- function() 1 - ((n - 1) / (n - p - 1)) * (1 - r2)
  wherry     <- function() 1 - ((n - 1) / (n - p)) * (1 - r2)
  olkin_pratt <- function() 1 - (((n - 3) * (1 - r2)) / (n - p - 1)) * Re(hypergeo::hypergeo(1, 1, (n - 3) / (n - p - 1), r2))
  olkin_pratt1 <- function() 1 - (((n - 3) * (1 - r2)) / (n - p - 1)) * (1 + (2 * (1 - r2)) / (n - p - 1))
  olkin_pratt2 <- function() 1 - (((n - 3) * (1 - r2)) / (n - p - 1)) * ((1 + (2 * (1 - r2)) / (n - p - 1)) *
    ((8 * (1 - r2) ^ 2) / ((n - p - 1) * (n - p + 3))))
  pratt      <- function() 1 - (((n - 3) * (1 - r2)) / (n - p - 1)) * (1 + (2 * (1 - r2)) / (n - p - 1))
  claudy     <- function() 1 - (((n - 4) * (1 - r2)) / (n - p - 1)) * (1 + (2 * (1 - r2)) / (n - p - 1))

  x <- vapply(adj, do.call, numeric(1), list(), envir = environment())
  names(x) <- adj
  x
}

r2 <- function(x, rho2, ...) {
  r2 <- summary(x)$r.squared
  n <- length(residuals(x))
  p <- length(x$coefficients) - 1
}
```

```
c(r2 = r2, mean_r2 = mean_r2(rho2, n, p), var_r2 = var_r2(rho2, n, p),
  adjusted_r2(r2, n, p, ...))
}
```

Har nu implementerat funktionerna ovan i `r2samplesize`-paketet och kan därmed framöver anropa dem därifrån.

4 2016-03-17

Jobbade med Teds projekt.

5 2016-03-18

Catched up on Pocket etc.

5.1 Läsning av (Croux and Dehon 2003)

Tolkar nämnaren i formeln för R^2 som residualvarians för modell med enbart intercept.

UNDersöker robustness av R^2 , dvs hur enskilda outliers påverkar. Finns föreslagna robusta skattningar av R^2 sedan tidigare som inte bygger på (men som kan fås till) till den vanliga LS-skattningen under vissa förhållanden.

I huvudsak väldigt teoretiskt och torr men ändå med det viktiga budskapet att den vanliga LS-skattningen inte är robust.

5.2 Läsning av (Renaud and Victoria-Feser 2010)

Handlar också om robusthet. Observera att vi tidigare sett att R^2 är ganska robust avseende ev skevhet i underliggande fördelning men vi är väl helt enkelt känsligare för outliers än för det.

Det finns många olika förslag på robusta estimators, även för små stickprovsstorlekar. DOck hade man hittills inte undersökt dem empiriskt med simuleringar etc, vilket sker i denna artikel.

Noterar att R^2 avser korrelation mellan y och bästa möjliga linjärbokobination av x (eftersom ju alla sådana linjärbokobinationer har samma korrelation).

It is important to note that normality is not assumed, merely the existence of the second moments. Compared to covariance penalty methods, although the R^2 is solely based on the covariance penalty, it plays an important role in model fit assessment. It should certainly not be used as a unique model fit assessor, but can provide a reasonable and rapid model fit indication.

The R^2 is usually presented as the quantity that estimates the percentage of variance of the response variable explained by its (linear) relationship with the explanatory variables.

Gör simuleringar med lite blandade data, kontinuerliga ovh dummy-variabler. Testar 8 olika estimat.

Man kan ju iofs också tänka att en utliggare är verklig och att korrelation bara ska mäta linjärt samband, vilket man kanske alltså inte har fullt ut ifall även utliggare ska kunna förklaras mha av modellen???

5.3 Läsning av (Asuero, Sayago, and González 2006)

Poängterar att korrelation och regression mäter olika saker.

For either correlation or for regression models, the same expressions are valid, although they differ significantly in meaning.

Nämner också att skattningen inte robust men att Spearmans korrelation då är ett alternativ som bara tar hänsyn till ranking och inte exakta värden.

F.ö. en mkt bra, grundläggande och översiktlig artikel som kan rekommenderas som en första introduktion till ämnet i största allmänhet!

The correlation coefficient, r , is a standardized index for which the value does not depend on the measurement scales of the variables. Its values lies (58) in the range $(-1, 1)$, and its squared value describes (59) the proportional reduction in variability of one variable when the other is held constant. It is worth noting that since for any r^2 other than 0 or 1, $|r^2| < r$; r may give the impression of a closer relationship (60) between x and y than does the corresponding r^2 . Therefore, although the correlation coefficient is by itself an important measure of relationship between the variables, it is R squared that permits comparison of the strengths of relationships. The reasons for making a distinction (41, 5) between r and R are that i) r is a measure of association between two random variables, whereas R is a measure between a random variable y and its prediction y from a regression model; ii) R is always well defined, regardless of whether the independent variable assumed to be random or fixed. In contrast, calculating the correlation between a random variable, y , and a fixed predictor variable x , that is, a variable that is not considered random, makes no sense.

Because r^2 gives the proportion of variance that is common between the two variables x and y , the uncommon or unique variance is the remainder, and this is known as the coefficient of nondetermination and is usually symbolized as $k^2 = 1 - r^2$. Some statisticians refer to (61) the coefficient of alienation k , which indicates the degree of lack of relationship.

5.4 Läsning av (Hössjer 2008)

OBS! Skriven av en Stockholmare! :-)

Handlar om mixed models. Kanske värt att nämna i ev samband med fixed och random models? Men känns i övrigt lite off topic vilket gör att jag inte läser mer än abstract.

5.5 Läsning av (Burton et al. 2006)

OBS! Handlar inte om ämnet utan om design av simulationsstudier inom medecin.

Föreslår att man vid simuleringsstudier, likväl som vid RCT etc upprättar ett studieprotokoll och att alla val av storlekar och metoder etc motiveras väl. Se fig 1 för uppställning av vad man bör tänka på. Man bör ange vilken slumpgenerator och med vilken seed som används. Måste tänka på hur olika variabler korrelerar, ofta orimligt att verkligen ha helt okorrelerade data. Bör om möjligt undersöka hela paramterspejset eller åtminstone den del som är av intresse vid tillämpning.

Olika mått för performance mearure kan och bör rapporteras, utöver ren bias även andel bias av hela måttet, standardiserad bias, MSE etc.

Genomgång har visat att väldigt många simuleringsstudier är undermåligt beskrivna. Skärpning!

Referenser

- Asuero, a. G., A. Sayago, and a. G. González. 2006. "The Correlation Coefficient: An Overview." *Critical Reviews in Analytical Chemistry* 36 (July): 41–59. doi:[10.1080/10408340500526766](#).
- Bobko, Philip. 2001. "A Review of the Correlation Coefficient and its Properties." In *Correlation and Regression: Applications for Industrial Organizational Psychology and Management.*, 12–42. doi:[10.4135/9781412983815](#).
- Burton, Andrea, Douglas G. Altman, Patrick Royston, and Roger L. Holder. 2006. "The design of simulation studies in medical statistics." *Statistics in Medicine* 25: 4279–92. doi:[10.1002/sim.2673](#).
- Cattin, Philippe. 1980. "Estimation of the Predictive Power of a Regression Model." *Journal of Applied Psychology* 65 (4): 407–14. doi:[10.1037//0021-9010.65.4.407](#).
- Claudy, J. G. 1978. "Multiple Regression and Validity Estimation in One Sample." *Applied Psychological Measurement* 2 (4): 595–607. doi:[10.1177/014662167800200414](#).
- Croux, Cristophe, and Catherine Dehon. 2003. "Estimators of the multiple correlation coefficient: Local robustness and confidence intervals." *Statistical Papers* 44 (3): 315–34. doi:[10.1007/s00362-003-0158-7](#).
- Ezekiel, and Mordecai. 1929. "The Application of the Theory of Error to Multiple and Curvilinear Correlation." *Journal of the American Statistical Association* 24 (165): 99–104.
- Hössjer, Ola. 2008. "On the coefficient of determination for mixed regression models." *Journal of Statistical Planning and Inference* 138 (10): 3022–38. doi:[10.1016/j.jspi.2007.11.010](#).
- Olkin, Ingram, and Jeremy D. Finn. 1995. "Correlations redux." *Psychological Bulletin* 118 (1): 155–64.
- Olkin, Ingram, and J.W. Pratt. 1958. "Unbiased estimation of certain correlation coefficients." *The Annals of Mathematical Statistics* 29 (1): 201–11. doi:[10.2307/2237306](#).
- Renaud, Olivier, and Maria Pia Victoria-Feser. 2010. "A robust coefficient of determination for regression." *Journal of Statistical Planning and Inference* 140 (7). Elsevier: 1852–62. doi:[10.1016/j.jspi.2010.01.008](#).
- Shieh, Gwown. 2010. "Estimation of the simple correlation coefficient." *Behavior Research Methods* 42 (4): 906–17. doi:[10.3758/BRM.42.4.906](#).
- Skidmore, Susan Troncso, and Bruce Thompson. 2011. "Choosing the Best Correction Formula for the Pearson r² Effect Size." *The Journal of Experimental Education* 79 (3): 257–78. doi:[10.1080/00220973.2010.484437](#).
- Wang, Zhongmiao, and Bruce Thompson. 2007. "Is the Pearson r² Biased, and if So, What Is the Best Correction Formula?" *The Journal of Experimental Education* 75 (2): 109–25. doi:[10.3200/JEXE.75.2.109-125](#).
- Wherry, R. 1931. "A new formula for predicting the shrinkage of the coefficient of multiple correlation." *The Annals of Mathematical Statistics* 2 (4): 440–57. [http://www.jstor.org/stable/2957681\\$/\\$backslash\\$npapers2:/publication/uuid/F3D4916B-BB98-4094-A459-DF4387AC9610](http://www.jstor.org/stable/2957681$/$backslash$npapers2:/publication/uuid/F3D4916B-BB98-4094-A459-DF4387AC9610).
- Wishart, Author J, T Kondo, and E M Elderton. 1931. "The Mean and Second Moment Coefficient of the Multiple Correlation Coefficient, in Samples from a Normal Population." *Biometrika* 22 (3/4): 353–76.
- Yin, Ping, and Xitao Fan. 2001. "Estimating R² Shrinkage in Multiple Regression: A Comparison of Different Analytical Methods." *The Journal of Experimental Education* 69 (2): 203–24. doi:[10.1080/00220970109600656](#)