Några funderingar om NPS

Bakgrund. NPS ('Net Promoter Score') finns beskrivet i ett speciellt SFK-dokument och är tänkt att användas för att redovisa ett resultat från en medlemsenkät. Nedanstående text är en fundering på några av måttets egenskaper. (Variationen hos NPS utreds i Appendix A.)

Utfallet från en medlemsenkät (på en heltalsskala (0-10)) delas in i tre proportioner – 'Ambassadörer' (p_1), 'Passiva' (p_2) samt 'Kritiska' (p_3). NPS definieras på följande sätt:

$$\text{NPS} = \frac{\text{antal ambassadörer} - \text{antal kritiker}}{\text{antal svarande}} \cdot 100$$

Om ett mätetals användning är begränsad på något sätt bör det anges ganska tydligt (om det inte är självklart). I läroböcker hittar man ofta dylika begränsningar typ 'för alla x > 0' eller för 'n > 100', osv. Något sådan information är svår att hitta när det gäller NPS. Nedanstående hypotetiska fall kan vara belysande (för enkelhetens skull har bägge fallen 100 svarande och alltså inget bortfall):

Fall	Antal ambassadörer	Antal passiva	Antal kritiker
A	3	94	3
B	48	4	48

En beräkning ger följande två NPS-värden:

$$\text{NPS} = \frac{3-3}{100} \cdot 100 = 0 \qquad \text{NPS} = \frac{48-48}{100} \cdot 100 = 0$$

Det är alltså märkligt att två helt skilda fall ger samma NPS-värde. Fall A kan omöjligtvis ge upphov till samma bekymmer och aktivitet som fall B där cirka hälften av alla svarande är kritiker.

Om man i stället sätter vikter eller kostnader på de olika kategorierna kan man uppnå åtminstone en viss diskrepans. Antag att vi har följande påhittade värden: 'Ambassadör 1', 'Passiva 0', 'Kritiker -10' där alltså en enda 'Kritiker' måste uppvägas av 10 'Ambassadörer':

$$\text{NPSkost} = \frac{3 \cdot 1 + 94 \cdot 0 + 3 \cdot (-10)}{100} \cdot 100 = -0.27$$

$$\text{NPSkost} = \frac{48 \cdot 1 + 4 \cdot 0 + 48 \cdot (-10)}{100} \cdot 100 = -4.32$$

Detta synsätt ger en klar differentiering mellan de två fallen. Observera dock att 'NPSkost' inte finns utan bara ett förslag, en av flera möjligheter.

Ytterligare analys. (Litteraturen anger en handfull krav på ett mätetal. Dessa krav är viktiga men är mer av matematisk natur så de utelämnas här.)
NPS är alltså en kvot mellan två heltal och definitionen kan, efter lite omarbete, också skrivas som skillnaden mellan två av ovanstående proportioner multiplicerat med faktorn 100:

$$\text{NPS} = \left(p_1 - p_3\right) \cdot 100$$

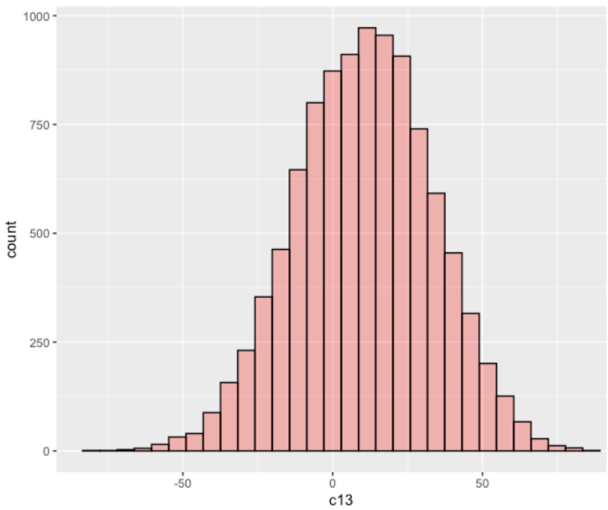
Resultatet avrundas till heltal och utfallet kan bli +/-100. (Notera att det inte är självklart att ta *skillnaden* mellan proportioner, man kan också tänka sig kvoten mellan dem dvs p_1/p_3 men det skulle ge skalan 0 till oändligheten.)

Om *samtliga* medlemmar svarade korrekt och enligt egen åsikt skulle vi exakt veta de olika proportionerna och sålunda handla därefter. Svarsprocent är dock ofta mycket lägre än 100% så resultaten måste anses vara stickprov ur populationen och därmed innehålla en slumpkomponent.

Simulering. Nedanstående simulering är baserad på att proportionen 'Ambassadörer' och 'Kritiska' är 0.50 respektive 0.40. Följande parametervärden och antagande är använda:

- Storlek på region/sektion: 50 medlemmar (mer än hälften av alla regioner/sektioner är mindre eller mycket mindre än 50 medlemmar)
- Svarsprocent: 35 %
- Proportion *Ambassadörer*: 50 %
- Proportion *Kritiska*: 40 %

Med ovanstående parametervärden simulerades 10000 enkätomgångar (två gånger). Resultatet visas i två grafer nedan - ett histogram över alla utfall (från omgång 1) samt ett histogram över möjliga förändringar mellan de två omgångarna.
Observera att *alla förutsättningar är konstanta*, den visade variationen består *enbart* av det slumpmässiga resultatet, inga förbättringar eller försämringar föreligger.



Figur 1. Simulerat utfall från 10000 beräknade NPS-värden med ovanstående parametervärden.

Kommentar figur 1. Histogrammet ovan består av ett antal distinkta staplar (detta är ju naturligt eftersom resultatet är en division av två heltal). Simuleringen består av 10000 enkätundersökningar (det kunde lika gärna varit 50000 eller 100000 simulerade undersökningar men utfallet och tolkningen hade inte blivit ännu klarare).

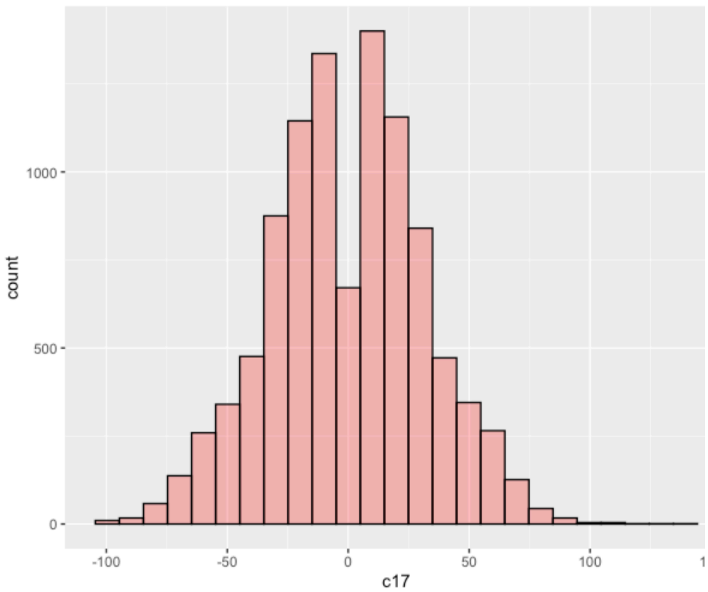
Det förväntade NPS-medelvärdet är 10 eftersom proportionen 'Ambassadörer' är 0.50 och proportionen 'Kritiska' är 0.40, detta stämmer också med histogrammet ovan. Den naturliga variation i NPS-värdet tycks ligga mellan, säg, -80 till +90.

Simulerat utfall (fig 1)

Variable	N	Mean	StDev	Minimum	Median	Maximum
NPS-värde avrundat	10000	10.356	22.166	-78.000	11.000	89.000

Skillnad mellan två mättillfällen. Det är ju helt naturligt att jämföra två olika mättillfällen och ställa frågan "Har vi blivit bättre eller sämre?"

Antag att vi har NPS-värde från två kvartal beskrivet som 'Undersökning kvartal Q1 2016' och ''Undersökning kvartal Q1 2017' och att vi beräknar förändringen mellan dessa två mätvärden. Följande histogram visar 10000 sådana jämförelser.



Figur 2. Skillnad mellan två omgångar av 10000 beräknade NPS-värden.

Kommentar figur 2. I medeltal tycks ju förändringen vara noll vilket är helt enligt förväntningarna ty det finns ju ingen förändring, alla resultat är utfall av slumpen. Trots det är det mycket stor risk (cirka 50%) att skillnaden mellan Q1 och Q2 felaktigt uppfattas som en försämring (kanske med medföljande negativa reaktioner). På samma sätt är det cirka 50% risk att skillnaden felaktigt uppfattas som en positiv framgång ("Jag visste att det gångna årets åtgärder skulle få en sådan positiv effekt!").

Simulerat utfall (fig 2)

Variable	N	Mean	StDev	Minimum	Median	Maximum
NPS-värde (skillnad)	10000	-0.038	31.708	-122.000	0.000	122.000

Ytterligare kommentar. I det 'vanliga' kvalitetsarbetet, då man arbetar med proportioner (ofta kallad felkvot) brukar man använda den beräknade kvoten och ett antal väl grundade metoder för analysen. Det finns då ingen anledning att hitta på nya mått eller göra det svårare än nödvändigt.

Bortfallet. Den proportion som inte svarar på en undersökning ger ofta upphov till problem. Hur skall den hanteras? Skall alla dessa kallas 'Ambassadörer' eller 'Kritiska'? Skall den observerade proportionen 'Kritiska' även appliceras på bortfallet och på så sätt artificiellt pressa upp stickprovsstorleken att bli samma som populationen och sålunda trolla bort den slumpmässiga inverkan?

En dyr KPI. Att uppfinna en KPI ('Key Performance Indicator') är en grannlaga uppgift (NPS är ju en KPI/mätetal/nyckeltal etc.), inte minst angående dess statistiska och användbara egenskaper.

På Ericsson-fabriken i Norrköping gjorde fabriksledningen en dyrköpt erfarenhet. Man skapade en KPI som skulle användas till ett bonussystem för de kollektivanställda. KPIen bestod av summan av tre delar – en kvalitetsdel (en enkel felprocent), en effektivdel (grundat på tidsutnyttjande) samt en kostnadsdel.

KPIen – som var resultatet av en förhandling mellan lönesättande kontor, företagsledning samt facket – hade ett övre gränsvärde. Om detta överskreds så betalades en bonus (kr/tim) till varje kollektivanställd. Om gränsvärdet inte överskreds så blev det 0 kr/tim. Hela upplägget var alltså enkelt att förstå och enkelt att administrera.

Det fanns dock ingen stresstestning eller simulering eller dylikt av konstruktionen och det blev så småningom tydligt att de slumpmässiga förändringarna ofta utlöste utbetalning av bonus utan att någon egentlig förbättring av kvalitet eller effektivitet kunde märkas.

Företaget ville så snart som möjligt omförhandla och 'höja ribban' men av naturliga skäl bromsade facket. Senare när Stockholm och högre nivåer på facket reagerade ändrades dock systemet.

Avslutningsvis. Det måste vara en naturlig och vanlig uppgift för en kvalitetsorganisation – och dess medlemmar – att ha en kritisk inställning till utformning, konstruktion, användning och redovisning av mätmetoder eller andra beslutskriterium, både hård- och mjukvara.

Appendix A (analys inklusive simuleringskod (Minitab och R i R-studio)

Antag att vi har följande definitioner:

p_1 = proportionen *Ambassadörer*

p_2 = proportionen *Passiva*

p_3 = proportionen *Kritiska*

Variationen $V(p_1)$ (dvs *variansen*) för en sådan proportion blir (n är stickprovets storlek):

$$\mathrm{V}(p_1) = \frac{p_1 \cdot (1 - p_1)}{n}$$

Ovan anges att NPS kan skrivas på följande sätt:

$$\mathrm{NPS} = (p_1 - p_3) \cdot 100$$

NPS är alltså en s.k. *linjärkombination* (dvs summa (eller skillnad) mellan mätvariabler) och variationen av NPS påverkas av variationen i dessa. Ibland finns det en samvariation (s.k. *kovarians*) mellan mätvärdena och denna kovarians adderas till den totala variansen. Kovariansen är antingen negativ, noll eller positiv, beroende på de aktuella mätvariablerna.

(Man har negativ kovarians om det ena mätvärdet tenderar att vara högre då det andra mätvärdet tenderar att vara lägre, i ett diagram skulle detta vara en nedåtlutande tendens. Negativ kovarians gör att den totala variationen minskar, vilket naturligtvis är önskvärt).

Kovariansen mellan $p_1 - p_3$ torde vara negativ ty om den ena proportionen är lite högre kan man förvänta sig att den andra är lite lägre (summan av de tre portionerna är ju alltid 1).

Denna kovarians $Cov(p_1, p_3)$ har följande matematiska uttryck:

$$\mathrm{Cov}(p_1, p_3) = -\frac{p_1 \cdot p_3}{n}$$

Det allmänna uttrycket för variansen för linjärkombinationen NPS blir (se litteraturen för detaljer):

$$\mathrm{V(NPS)} = \left[\mathrm{V}(p_1) + \mathrm{V}(p_3) - 2 \cdot \mathrm{Cov}(p_1, p_3)\right] \cdot 100^2$$

Med insättning erhålles alltså följande:

$$\mathrm{V(NPS)} = \left[\frac{p_1 \cdot (1 - p_1)}{n} + \frac{p_3 \cdot (1 - p_3)}{n} + 2 \cdot \left[-\frac{p_1 \cdot p_3}{n}\right]\right] \cdot 100^2$$

Efter förenkling blir det till slut:

$$\mathrm{V(NPS)} = \left[\frac{p_1 + p_3 - (p_1 - p_3)^2}{n}\right] \cdot 100^2$$

För tolkningen av resultatet är det dock oftast bättre med standardavvikelsen *sigma* (som ju är roten ur variansen) och med indata från sidan 2 blir det följande (NPS-enheter):

$$\sigma_{\text{NPS}} = \sqrt{\left[\frac{0.5+0.4-(0.5-0.4)^2}{18}\right] \cdot 100^2} = 22.2$$

(Nämnare '18' kommer av en population om 50 där endast andelen 0.35 svarat). En vanlig tumregel är att +/- 3 sigma omfattar den slumpmässiga variationen, här cirka +/- 66. Vi förväntade oss ett NPS-värde på (0.5 – 0.4) • 100 dvs 10 men det är tydligt att den slumpmässiga variationen kan ge oss ett utfall långt från det förväntade och sålunda få oss att dra fel slutsatser.

Skillnad mellan två mätningar. I simuleringen (histogram 2) redovisas skillnaden mellan två mätningar med exakt samma förutsättningar (indata enligt sidan 2). En dylik skillnad är i analytiskt hänseende ytterligare en linjärkombination:

$$\text{Diff}_{\text{NPS}} = \text{Q1}_{\text{NPS}} - \text{Q2}_{\text{NPS}}$$

Standardavvikelsen för denna differens blir med ovanstående teoretiska ramverk:

$$\sigma_{\text{diff}} = \sqrt{22.2^2 + 22.2^2} = 31.4$$

Slutsats. Med gruppstorlek om n = 50 där 35 procent svarat blir skillnaden mellan två mättillfällen ganska varierande (figure 2 ovan) även om det inte finns *någon reell skillnad mellan kvartal 1 och 2*!!

En slutsats att skillnaden i verkligheten är positiv (eller negativ) är alltså helt fel och kan ge oanade konsekvenser. Mätmetoden kräver alltså ganska stora stickprov för att vara av något som helst värde.

Teoretisk standardavvikelse för skillanden blev ovan 31.4. Tre simulering gav värdena 31.6, 31.6, 31.9. Detta visar att den teoretiska härledningen är korrekt.

Simuleringskod (Minitab)

```
erase c1-c4000
erase k1-k1000

name c1 'Amb/Pass/Krit' c2 'Prop Amb/Pass/Krit'
name c3 'Slump' c4 'Enkät nr'

set c1
1 2 3
end

set c2
0.50 0.10 0.40          # Proportion 'Amb', 'Pass', 'Krit'
end


let k2 = 0.35           # Andel svarande.
let k3 = 50             # Antal medlemmar i en sektion/region.
let k4 = round(k2*k3)   # Antal svarande (stickprovets storlek).
let k5 = 10000          # Antal utskick av enkäter.

let k6 = k4*k5

Random k6 c3;
  Discrete C1 C2.

Set c4                  # K5 enkäter om k4 svarande.
(1:k5)k4                # Kolumn c4 är 'enkätnr'.
end

Let c6 = IF(c3=1;1;0)   # Alla 'Amb' blir 1:a, övr 0:a.


Name c7 "Prop Amb"

Statistics C6;          # Beräknar andelen 'Amb' per grupp
  By C4;                # och lägger resultatet i 'Prop Amb'.
  Mean 'Prop Amb'.

Let c9 = IF(c3=3;1;0)

Name c10 "Prop Krit"

Statistics C9;
  By C4;
  Mean 'Prop Krit'.

Let c12 = ('Prop Amb' - 'Prop Krit')*100

name c13 'NPS-värde avrundat'

let c13 = round(c12)     # NPS som heltal.

hist c13;
title 'Simulerad fördelning av NPS';
grid 2;
scale 1;
hdispl 0 0 0 0;
scale 2;
hdispl 0 0 0 0;
```

```
ldispl 0 0 0 0;
axlabel 1 'NPS-värde';
axlabel 2 ''.


# ----------------------------------

# Ovanstående data lagras i kolumn c15 och
# därefter köras samma simulering och lagras i c16.
# Skillnaden beräknas och lagras i c17:

let c17 = c15 - c16


hist c17;
title 'Simulerad förändring mellan två mätningar';
grid 2;
scale 1;
hdispl 0 0 0 0;
scale 2;
hdispl 0 0 0 0;
ldispl 0 0 0 0;
axlabel 1 'Skillnad i NPS-värde';
axlabel 2 ’'.
```

Simuleringskod (R i R-studio)

```
# NPS

k2 <- 0.35              # Andel svarande
k3 <- 50                # Antal medlemmar i gruppen
k4 <- round(k2*k3, 0)   # Antal svarande
k5 <- 10000             # Antal grupper
k6 <- k4*k5             # Totalt antal som svarat (1 eller 2 eller 3).

c3 <- sample(1:3, size = k6, replace = TRUE, prob=c(0.50, 0.10, 0.40))   # De 3
olika svarsalernativen

c6 <- ifelse(c3 == 1, 1, 0)     # 1 = "Ambassadör"
c9 <- ifelse(c3 == 3, 1, 0)     # 1 = "Kritiska"

c4 <- rep(1:k5, each = k4)      # Skapar gruppnummer

c7 <- aggregate(c6, list(c4), FUN = mean)   # Beräknar gruppmedelvärde
c10 <- aggregate(c9, list(c4), FUN = mean)

c12 <- (c7$x - c10$x)*100       # NPS-värdet dvs diff 'Ambassdörer' o 'Kritiska'

c13 <- round(c12, 0)


minaadata1 <- data.frame(c13)

library(ggplot2)

hist1 <- ggplot(minaadata1, aes(c13)) + geom_histogram(bins=30, color='black',
fill='red', alpha = 0.3, position = 'identity')
hist1
```

```
# -------------------------------------------------------- Nedan upprepas
simuleringen, resultat c14.

c3 <- sample(1:3, size = k6, replace = TRUE, prob=c(0.50, 0.10, 0.40))   # De 3
olika svarsalernativen

c6 <- ifelse(c3 == 1, 1, 0)     # 1 = "Ambassadör"
c9 <- ifelse(c3 == 3, 1, 0)     # 1 = "Kritiska"

c4 <- rep(1:k5, each = k4)      # Skapar gruppnummer

c7 <- aggregate(c6, list(c4), FUN = mean)    # Beräknar gruppmedelvärde
c10 <- aggregate(c9, list(c4), FUN = mean)

c12 <- (c7$x - c10$x)*100       # NPS-värdet dvs diff 'Ambassdörer' o 'Kritiska'

c14 <- round(c12, 0)

c17 <- c13 - c14

minaadata2 <- data.frame(c17)

library(ggplot2)

hist2 <- ggplot(minaadata2, aes(c17)) + geom_histogram(bins=25, color='black',
fill='red', alpha = 0.3, position = 'identity')
hist2

sd(c17)     # Standardavvikelse (att jämföra med teoretisk beräkning)
```