



Rapport

23. november 2018
Proj.nr. 2004280
Version 1
EVO/MT

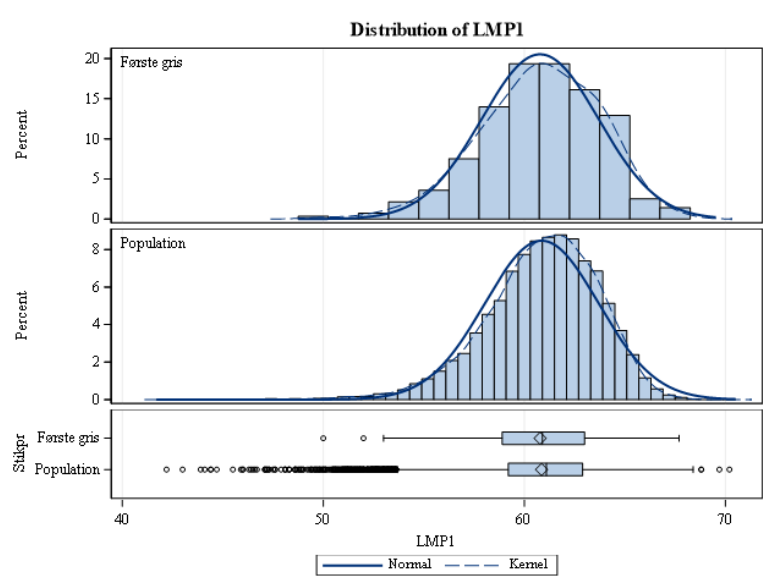
Principper for og forslag til repræsentative stikprøveplaner til analyse af konsekvensen af produktionsændringer for værdi- og kvalitetsvurdering

Eli V. Olsen

- Introduktion* Hovedparten af alle undersøgelser og analyser af populationer er baseret på stikprøver. Men hvis stikprøven ikke afspejler det fænomen, man ønsker at undersøge, kan der naturligvis ikke drages de rigtige konklusioner.
- Problemstilling* Mange undersøgelser vedrørende slagtekvalitet i bred forstand baseres på stikprøver udtaget på en slagtelinje. Undersøgelserne omfatter ofte dyre målinger, hvorfor stikprøverne typisk er små repræsentative stikprøver. Repræsentativitet opnås ved tilfældig udvælgelse. Spørgsmålet er blot, hvordan man sikrer, at udvælgelsen reelt er tilfældig og repræsentativ. Problemet er ikke blevet mindre i takt med, at leverancerne er blevet større. Hvis en stikprøve på 20 slagtekroppe udvælges ved at udtage hver 10. slagtekrop, kan man risikere, at alle slagtekroppe kommer fra den samme leverance, som ikke nødvendigvis repræsenterer alle leverancer.
- Formål* Formålet med denne rapport er at vise eksempler på forskellige udvælgelsesstrategier for stikprøver og diskutere deres egnethed. Stikprøvernes repræsentativitet vurderes ved at sammenligne undersøgelsesresultatet med det resultat, man ville opnå, hvis hele "populationen" blev undersøgt.
- Datagrundlag* Nedenstående eksempler er baseret på målt kødprocent og vægt samt de veterinære bedømmelser af forekomst af betændelse, knogleproblemer samt bylder og sår. Disse oplysninger findes for hele "populationen", som i dette tilfælde er en uges leverancer til et mindre slagteri (ca. 5.000 slagtninger/dag).
- Eksempel 1* *Kan én slagtekrop fra hver leverance repræsentere hele populationen?*
- Dataeksemplet indeholder data fra 276 leverancer af i alt 23.974 grise. Fordelingen af kød% for stikprøven bestående af den første slagtekrop i hver leverance, der måles på slagtelinjen, er vist sammen med fordelingen af kød% for alle slagtekroppe, se figur 1.

De to fordelinger er sammenlignet ved hjælp af en t-test, hvor der kan tages højde for antalsforskellen. De estimerede standardafvigelser for stikprøven og hele populationen er stort set ens ($p=58\%$). De estimerede kød%-gennemsnit er ligeledes ens: 60,8 hhv. 60,9 ($p=65\%$).

De estimerede standardafvigelser for afregningsvægten er signifikant forskellige (9,4 kg hhv. 7,0 kg, $p<0,001$), men gennemsnitsvægtene kan antages at være ens (87,7 kg og 87,1 kg, $p\approx 0,2$).



Figur 1. Fordeling af kød% i stikprøve og i hele populationen.

Hvis stikprøveudvælgelsen udelukkende baseres på leverancer med flere end 100 grise, bliver resultaterne som anført i tabel 1.

Tabel 1. Store leverancer: Stikprøve med én gris pr. leverance vs. hele populationen.

Stikprøve	N	Gns kød%	Std Dev	Std Err
Første gris	85	61,3	2,63	0,29
Population	15794	61,1	2,71	0,02
Diff. (1-2)		0,2	2,71	0,29

Kommentar

Selv om t-testen underbygger en antagelse om, at stikprøvens gennemsnit svarer til hele populationens gennemsnit, og tager højde for antalsforskellen, giver resultatet ikke nødvendigvis mening. En stikprøve med 85 målinger, som i tabel 1, giver en betydeligt dårligere bestemmelse (std err=0,29) af gennemsnittet, end når hele populationen er til rådighed (std err=0,02). Det betyder, at gennemsnittet i en ny stikprøve kan afvige op til $\pm 2 \times 0,29$ fra "sandheden" (populationsgennemsnittet) alene på grund af tilfældigheder.

Ved kalibrering af klassificeringsudstyr er det vigtigt, at stikprøven både repræsenterer populationens gennemsnit og spredning, hvilket stort set er tilfældet i dette eksempel. MEN stikprøven er alligevel ikke repræsentativ. Stikprøven, som består af den første gris i alle de store leverancer, viser sig at være sogrise! Fra tidligere undersøgelser vides, at so- og galtgrise har forskelligt kødindhold. I denne population har sogrisene næsten 2 kød%-enheder mere kød end galtgrisene (62,1 hhv. 60,2), så stikprøvens sogrise er ikke engang repræsentative for populationens sogrise.

Den valgte strategi er altså ikke anbefalelsesværdig til kalibrering af klassificeringsudstyr.

Eksempel 2 *Hvor meget varierer kød%’en fra dag til dag?*

I dette eksempel udgør den gennemsnitlige kød% for en leverance en observation. Observationerne opdeles svarende til store (>100 grise) henholdsvis små leverancer. De i alt 279 leverancer over fem dage er fordelt som vist i tabel 2.

Tabel 2. Fordeling af leverancer mht. størrelse og slagtedag

Leverance	SLAGTEDAG					
	20180621	20180622	20180625	20180626	20180627	Total
Fre- quency						
Lille	42 72%	35 66%	37 65%	33 63%	44 79%	191 69%
Stor	16 28%	18 34%	20 35%	19 37%	12 21%	85 31%
Total	60 21%	58 19%	59 21%	55 19%	58 20%	276 100%

Data er analyseret med en tosidet variansanalyse med gennemsnitlig kød% pr. leverance som afhængig variabel, og leverancestørrelse og ugedag som faktorer. Der er ingen signifikant vekselvirkning mellem faktorerne. De store leverancer har en gennemsnitlig kød% på 61,0 (std err=0,16), og de små har en gennemsnitlig kød% på 60,3 (std err=0,11). Forskellen er signifikant ($p < 0,001$).

Der er også en signifikant variation mellem dage med en maksimal forskel på op til 1 kød%, se tabel 3.

Tabel 3. Ugevariation - tal med samme bogstav er ikke signifikant forskellige.

SLAGTEDAG	Gns. kød%	Std Err
20180621	60,6	0,20
20180622	61,1 ^a	0,20
20180625	60,3 ^b	0,20
20180626	60,8 ^a	0,21
20180627	60,2 ^b	0,21

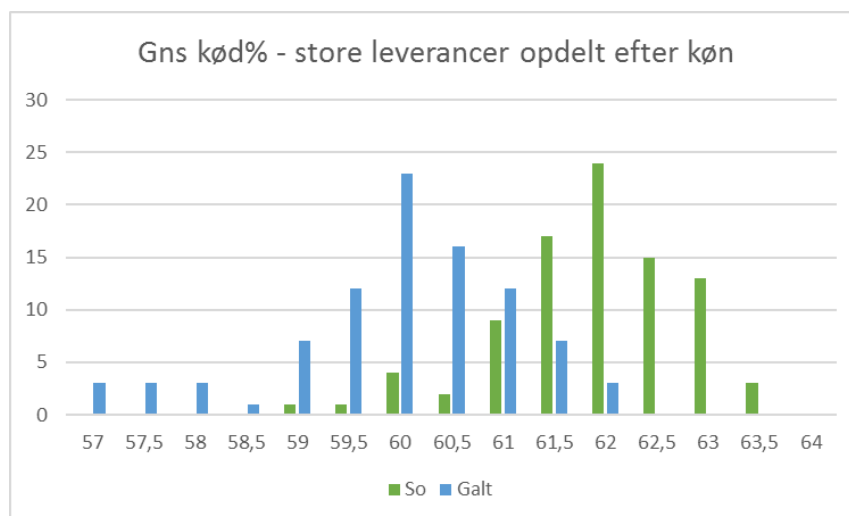
Kommentar Opdelingen i store og små leverancer medfører, at observationerne i princippet er bestemt med forskellig præcision. Præcisionen for observationer i gruppen med store leverancer med flere end 100 grise er $\frac{1}{\sqrt{100}}$ x standardafvigelsen eller bedre, medens observationer svarende til små leverancer er ringere. Da det primære formål er at analysere dagsvariationen, og da observationerne er nogenlunde ligeligt fordelt på faktorerne, se tabel 2, kan forskel i de enkelte observationers præcision ignoreres. Den anførte std. err i tabel 3 refererer til variationen mellem leverancegennemsnit.

Forskellen mellem dage er betydelig, hvis opgaven er at udvælge en stikprøve til kalibrering af klassificeringsudstyr. Det er derfor vigtigt at udvælge over flere dage.

Eksempel 3 *Hvor meget varierer kød%’en mellem producenter og køn?*

Figur 2 viser den gennemsnitlige kød% for hver leverance opdelt efter køn. Leverancer af sogrise har i gennemsnit 62% kød, men varierer fra 59 til 64% (std.afv.=0,9), medens galtgriseleverancer i gennemsnit har 60% kød, varierende fra 57 til 62% (std.afv.=1,1). Forskellen i leverancerne kan skyldes racer og krydsninger, men også management og valg af foder.

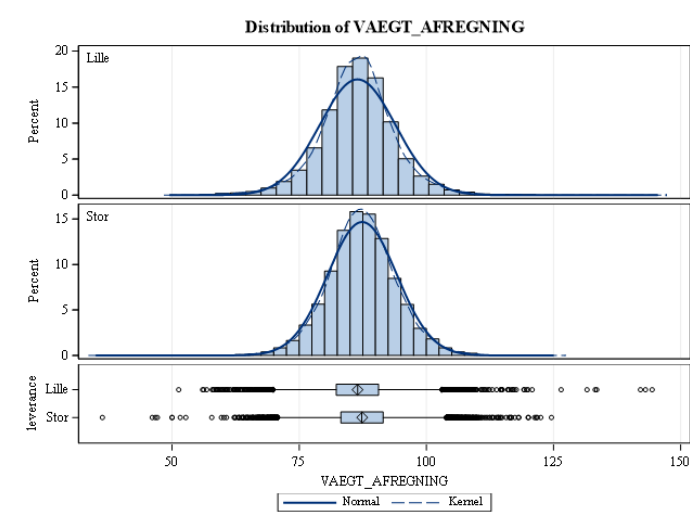
Kommentar Det har derfor stor betydning, at stikprøven til kalibrering af klassificeringsudstyr har lige mange so- og galtgrise fra forskellige leverandører. Hvis dette ignoreres, vil man kunne risikere at udtage en sogristikprøve fra en leverance med gennemsnitligt 60% kød og en galtgrise-stikprøve fra en leverance med gennemsnitligt 62% kød. Man vil opnå den forventede køn-forskel, blot med modsat fortegn!



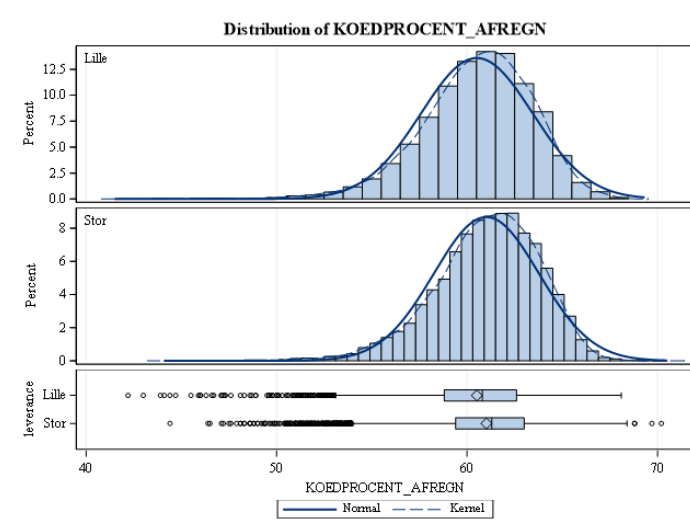
Figur 2. Fordeling af gns. kød% per leverance og køn.

Populationsbeskrivelse

Hvis slagtekroppene opdeles efter, om de kommer fra en stor leverance, som her er defineret som leverancer med flere end 50 so- og 50 galtgrise, så vil 70% af slagtekroppene komme fra store leverancer. Slagtekroppe fra store leverancer er i gennemsnit næsten 1 kg tungere og har 0,5% mere kød end slagtekroppe fra små leverancer.



Figur 3. Fordeling af afregningsvægt opdelt efter leverancestørrelse.



Figur 4. Fordeling af kød% opdelt efter leverancestørrelse.

Diskussion og konklusion

Ovenstående eksempler viser, at der er systematisk variation mellem dage, køn og leverancestørrelse med hensyn til kød%. Dertil er der en tilfældig variation indenfor hver gruppering. Dette afspejler en variation bestemt af forældredyrene (race/krydsning, avlsindeks), fodersammensætning og formentlig også staldforhold og management. Ved kalibrering af klassificeringsudstyr er det vigtigt, at der opnås en prædiction af kødindholdet, som er en gennemsnitsværdi for alle typer slagtekroppe, med de samme målte egenskaber.

Grisene slagtes ikke i tilfældig orden, men i grupper svarende til leverancerne. Derfor er det nødvendigt at anvende stratifikation ved stikprøveudvælgelsen. Det vil sige, at stikprøven må udtages fra delgrupper (strata), der er forskellige med hensyn til kød%.

Eksemplerne ovenfor viser, at stikprøven skal have lige mange so- og galtgrise udvalgt fra forskellige leverandører. 70% af slagtekroppene skal komme fra store besætninger, resten fra små.

Repræsentation af variationen mellem leverancer opnås ved at udtage stikprøven over flere dage, idet dagsvariationen formentlig primært afspejler forskellen mellem leverancerne.

Endelig skal slagtekroppene udvælges tilfældigt fra en leverance og ikke systematisk, som for eksempel den første i hver leverance.

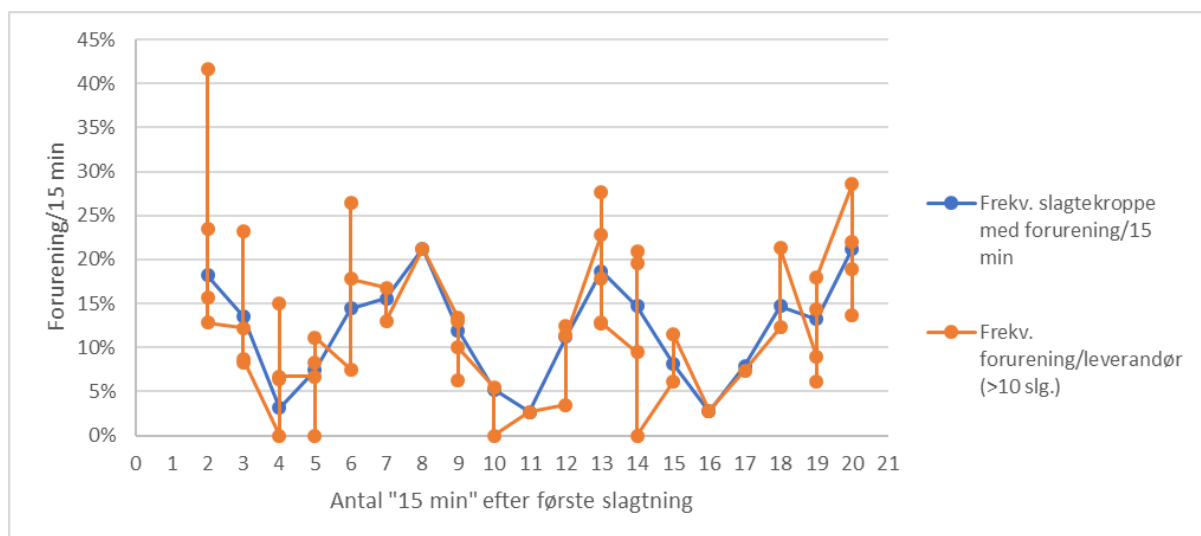
Udvælgelse af meget små stikprøver

Eksempel

Udvælgelse af små stikprøver

Mikrobiologiske test er typisk baseret på små stikprøver på grund af de store omkostninger, der er forbundet med analyserne. Hvis undersøgelsen fx går ud på at undersøge slagtekropsforureninger, viser figur 5, hvor vigtigt det er at vælge stikprøven over et længere tidsforløb. De blå punkter er gennemsnitlig forureningsfrekvens (hyppighed af slagtekroppe med forurening) i 15 minutters intervaller, medens de orange punkter er gennemsnitlig forureningshyppighed opdelt efter leverance i de samme 15 minutter.

På den pågældende slagtelinje slagtes i gennemsnit 150 grise pr. 15 min. Udtages stikprøven i tidsrummet markeret som "8" på figuren, vil man udtage fra én leverance, med relativ høj forureningsfrekvens (20%), medens man udtager fra én leverance, med relativ lav forureningsfrekvens (<5%), hvis der kun udtages prøver i tidsrummene "11" eller "16".



Figur 5. Eksempel på forureningsfrekvenser over en produktionsdag.

Diskussion og konklusion Slagtekropsforurening er ikke nødvendigvis relateret til leverandør. Der er muligvis en interaktion mellem leverancer og håndteringen på slagte- linjen. Hvis producenten sørger for, at maveindholdet er mindst muligt, er risikoen for forurening ved udtagning af fedtende og tarmsæt for- mentlig mindre.

Udtagning af stikprøve til fastlæggelse af mikrobiologisk status på slag- telinjen skal derfor foretages over så lang en periode som muligt. Men det vil i alle tilfælde være vanskeligt at opnå et repræsentativt gennem- snit med stikprøvestørrelser under 10 eller 20 prøver.

Kreditering Rapporten er udarbejdet som resultat af arbejdet i arbejdspakken *WP4. Metoder til validering af måledata (2016-2018)* i hovedprojektet *Udnyt- telse af detaljeret råvareviden*. Projektet er finansieret af Svineafgifts- fonden, 2018.

Deltagere i arbejdspakken:

Eli Vibeke Olsen (hovedforfatter til rapporten),

Dennis Brandborg Nielsen, Peter Vorup, Morten Askjær Hass.