

Træningsspørgsmål efter 9. forelæsning/øvelse

Der skal afholdes en konkurrence mellem to basketballhold, fra to konkurrerende skoler, A og B.

Det skal foregå på den måde, at hver spiller har “*sin egen kurv*” og løbende bliver forsynet med bolde. Hver spiller kaster indtil vedkommende har fået bolden i kurven 5 gange, og den benyttede tid til dette registreres.

1. Er det fornuftigt at sammenfatte holdets resultat ved hjælp af den gennemsnitlige tid, der bruges på denne opgave?
 - (a) Nej, fordi fordelingen formentlig er ret skæv
Tja, det er svært at sige på forhånd, hvordan fordelingen vil se ud....
 - (b) Ja, hvis der er mange deltagere på hvert hold **Forkert**: P-værdien for en sammenligning af middelværdier er robust overfor skævheder o.l., når antallet er stort, men selve gennemsnittet får ikke nødvendigvis mere mening af den grund, med mindre der er tale om en fordeling med enkelte outliers.
 - (c) Ja, for fordelingen må forventes at være rimeligt symmetrisk
Tja, det er svært at sige på forhånd, hvordan fordelingen vil se ud....
 - (d) Nej, det er bedre at bruge medianen
Tja, statistisk set er det mere robust overfor mærkelige fordelinger, men man får næppe overbevist sportsverdenen om det, tænker jeg
 - (e) Ja, fordi det er sådan, man gør normalt
Korrekt: Hvis reglerne er sådan, så er det sådan, man gør.
2. Hvordan kan vi afgøre, hvilket af de to hold, der er bedst?
 - (a) Vi kan lave et uparret T-test til sammenligning af de to fordelinger
Delvist korrekt: P-værdien svarende til et uparret T-test er ret robust overfor skævheder, når antallet er stort.

- (b) Det må jo være det hold, der har brugt mindst tid på opgaven
Forkert: En lille forskel kan nemt skyldes tilfældigheder og behøver derfor ikke, at betyde, at det ene hold er bedre end det andet. Men hvis man vil kåre en vinder, så er det selvfølgelig tilstrækkeligt at have den korteste tid.
- (c) Man kan lave et nonparametrisk (Mann-Whitney) test til sammenligning af de to hold.
Korrekt: Det kan vi i hvert fald tillade os.
- (d) Det kan vi ikke, for der kan være så mange tilfældigheder, der indvirker på resultatet
Korrekt: Men det er jo det, vi undersøger ved at lave et test af identiske fordelinger.

Man beslutter at udvide konkurrencen for at nedbringe tilfældigheder ved udpegning af en vinder. Derfor skal spillerne nu alle have scoret 20 gange, inden man registrerer tidsforbruget, og man vil gå videnskabeligt til værks, så det kan ende uafgjort, hvis man ikke kan påvise forskelle ud over tilfældighedernes spil.

3. To spillere fra hold A når ikke at få bolden i kurven de 20 gange inden konkurrencen bliver afblæst. Hvordan skal vi håndtere dette?
 - (a) Vi sætter tidsforbruget for disse to spillere til den totale tid for konkurrencen og analyserer data ligesom ovenfor.
Forkert: Dette vil underestimere tidsforbruget for hold A, da tidsforbruget for disse spillere vides at *overstige* denne værdi - men ikke hvor meget....
 - (b) Baseret på det antal gange, de pågældende spillere scorede indenfor tidsfristen, kan vi estimere hvor lang tid, de ville bruge på at få alle 20 bolde i kurven, og dette estimat benytter vi til at analysere data som ovenfor.
Delvist korrekt: Dette er et forsøg på *imputation*, hvor man forsøger at erstatte en manglende værdi med et kvalificeret gæt.
 - (c) Vi kan ikke mere regne med gennemsnit og laver test for identiske medianer i stedet for.
Delvist korrekt: Dette kan godt fungere, fordi de to manglende værdier vides at være de højeste, og medianerne kan derved udregnes.

- (d) Vi behandler tiden for disse to spillere som censurerede observationer og sammenligner holdene med et log-rank test
Korrekt: Når vi behandler dem som censurerede, siger vi netop, at vi kun ved, at tidsforbruget er *større* end den totale tid for konkurrencen.
 - (e) Vi udelader disse to spillere af analysen, da de ikke har valide data
Forkert: Dette vil underestimere tidsforbruget for hold A, da tidsforbruget for disse spillere vides at være de allerstørste
4. En spiller fra hold B kommer for sent og får derfor reduceret sin tid til opgaven. Hvordan håndterer vi det?
- (a) Vi udelader denne spiller af analysen, da han ikke har valide data
Forkert: Det vil være unødigt spild af information, og det kunne føre til bias, hvis forsinkelsen kunne hænge sammen med manglende interesse (eller dygtighed) til basketball.
 - (b) Vi sætter hans tid til gennemsnittet for resten af holdet.
Forkert: Som ovenfor, og en klar suboptimal løsning, når han dog har været med noget af tiden.
 - (c) Vi medtager denne spiller i analysen med forsinket indgang.
Forkert: Den relevante tidsskala må her være det tidspunkt, den enkelte spiller starter med at forsøge at score. Der er derfor tale om en censureret observation, fordi han ikke når at deltage i lige så lang tid som de øvrige.
 - (d) Vi estimerer hans tid ud fra hvor mange gange, han scorede indenfor den tid, han faktisk havde til opgaven
Delvist korrekt: Dette er et forsøg på *imputation*, hvor man forsøger at erstatte en ufuldstændig observation med et kvalificeret gæt.
 - (e) Vi benytter hans faktisk anvendte tid som en censureret observation i en overlevelsesanalyse.
Korrekt: Når vi behandler den som censureret, siger vi netop, at vi kun ved, at tidsforbruget er *større* end den tid, han nåede at være med i konkurrencen.
5. En anden spiller fra hold B bliver ramt af en vildfaren bold og må udgå af konkurrencen ca. halvvejs. Hvordan håndterer vi det?

- (a) Vi estimerer hans tid ud fra hvor mange gange, han scorede indenfor den tid, han faktisk havde til opgaven
Delvist korrekt: Dette er et forsøg på *imputation*, hvor man forsøger at erstatte en ufuldstændig observation med et kvalificeret gæt.
 - (b) Vi udelader denne spiller af analysen, da han ikke har valide data
Forkert: Det vil være unødigt spild af information.
 - (c) Vi medtager denne spiller med en censureret tid.
Korrekt: Når vi behandler den som censureret, siger vi netop, at vi kun ved, at tidsforbruget er *større* end den tid, han nåede at være med i konkurrencen.
 - (d) Vi sætter hans tid til gennemsnittet for resten af holdet.
Forkert: Som ovenfor, og en klar suboptimal løsning, når han dog har været med noget af tiden.
6. En spiller fra hold A har virkelig problemer med sit boldøje og kan slet ikke ramme kurven. Halvvejs igennem turneringen gider han derfor ikke mere og forlader banen. Hvordan håndterer vi det?
- (a) Vi medtager denne spiller med en censureret tid, svarende til den tid, han faktisk deltog i konkurrencen.
Forkert: Der er information i det faktum, at han udgår fordi han ikke dur til det. Der ville altså være tale om informativ censurering, som giver bias i analysen.
 - (b) Vi udelader denne spiller af analysen, da han ikke har valide data
Forkert: Dette vil underestimere tidsforbruget for hold A, da tidsforbruget for denne spiller formentlig vil være *virkelig stort*.
 - (c) Vi har et svært problem, da der er tale om informativ censurering
Korrekt: Det er et meget ubehageligt problem, med kun få og vanskelige løsningsmuligheder.
 - (d) Vi sætter hans tid til et meget stort tal, da han må formodes at være meget dårlig til det
Tja...., men hvilket stort tal?
 - (e) Vi sætter hans tid til gennemsnittet for resten af holdet.
Forkert: Dette vil underestimere tidsforbruget for hold A, da tidsforbruget for denne spiller formentlig vil være *virkelig stort*.

7. Efter at konkurrencen er slut, og hold B er udpeget som vinder, kommer der klager fra hold A, fordi de mener, at deres spillere var væsentligt lavere end hold B, og at deres kurve derfor burde være sat noget lavere. Hvad skal man sige/gøre ved denne klage?
- (a) Den skal afvises, da man bør være enige om reglerne inden start.
Korrekt: Som konkurrence betragtet må dette være det fornuftige svar.
 - (b) Spillernes højde bør medtages som kovariat i analysen, så der bliver korigeret for den påståede højdeforskel.
Tja, det *kunne* man gøre, men det burde så være noget, man havde meldt ud på forhånd
 - (c) Man nedsætter hold A's tidsforbrug proportionalt med deres lavere højde
Forkert: Det er nok lige lovlig *ad hoc* blot at gange en mere eller mindre tilfældig faktor på....
8. Ville dit svar på spørgsmål 7 ændre sig, hvis formålet med konkurrencen ikke var at udpege en vinder, men at afgøre, om de to forskellige træningsformer, der benyttes på hold A og B, kunne påvises at have forskellig kvalitet?

Ja, for hvis der er en betydelig forskel på højderne på de to skoler, så vil sammenligningen være udtryk for en sammenblanding af effekten af højde og effekten af træningsformen. For at få *den rene* effekt af træning, ville vi derfor korrigere for højden ved at inddrage den som kovariat.

Og i så fald, ville det så give nogen mening at inddrage en interaktion mellem hold og højde?

Nej, det er vanskeligt at se, hvad der skulle berettige at inddrage en interaktion... Det skulle være noget med, at den ene træningsform virkede bedre for de høje, og den anden bedre for de lave...??