# Eksamen 2023 / 2024 - Sandsynlighed og statistik

Optimering af en Formel 1 racer

Af Jesper Bertelsen

Det amerikanske formel 1 hold MoneyGram Haas F1 Team har ansat to kørere – nr. 20 Kevin Magnussen og nr. 27 Nico Hülkenberg. I sæsonen 2023 har begge kørere haft alvorlige problemer med for meget dækslid i løbene. Under tidtagningerne har Hülkenberg generelt klaret sig bedre end Magnussen på en enkelt omgang med friske dæk, mens forskellene har udjævnet sig i løbene. Man mener, at forskellene på kørernes performance skyldes en uheldig kombination af kørernes individuelle kørestil, racerens opsætning og dens bremser.

MoneyGram Haas benytter bremser af det italienske mærke Brembo. Man ønsker at undersøge, om et andet fabrikat kan forbedre performance for Magnussen, uden at forringe den for Hülken- berg. Man ønsker af praktiske grunde at benytte den samme leverandør til begge biler. Navnet på det alternative bremsefabrikat er holdt hemmeligt. Vi kalder det X Breaking Systems (XBS).

De følgende opgaver handler om (konstruerede) data fra eksperimenterne.

Indholdsfortegnelse

[Eksamen 2023 / 2024 - Sandsynlighed og statistik 1](#_Toc155778367)

[Opgave 1. Bremseevne 2](#_Toc155778368)

[b. Lav en variansanalyse (ANOVA) på signifikansniveau 5 %, der belyser, om der mht. bremseevne er signifikant effekt af hver af de tre faktorer. Er der signifikante 2-faktor og 3-faktor interaktioner? Kom tilbage hertil. 3](#_Toc155778369)

[Opgave 2. Bremsernes kølesystem 4](#_Toc155778370)

[c. Opstil formlen for teststørrelsen for en Goodness of Fit test og oplys teststørrelsens fordeling og antal frihedsgrader. Kom tilbage hertil 5](#_Toc155778371)

[Opgave 3. Omgangstider 7](#_Toc155778372)

## Opgave 1. Bremseevne

MoneyGram Haas F1 Team tester de to bremsefabrikater ved at lade hver kører skiftevis køre igennem det samme sving på tid i deres normale racerbil. Før svinget er racerbilens fart indstillet til 300.0 km/t, og det kræver en nedbremsning til cirka 70 km/t, for at komme optimalt igennem svinget. For at vurdere bremsernes egenskaber under forskellige forhold køres der med tre forskellige typer af tørvejrdæk fra Pirelli.

Eksperimentet har således følgende 3 faktorer:

* Faktor A. Bremsefabrikat: 1: Brembo, 2: X Breaking Systems
* Faktor B. Racerkører: 20: Kevin Magnussen, 27: Nico Hülkenberg
* Faktor C. Dæktype: 1: Hard, 2: Medium, 3: Soft

Et billede, der indeholder tekst, skærmbillede, nummer/tal, Font/skrifttype

Automatisk genereret beskrivelseFor hver kombination af de 3 faktorer er der 5 gentagelser. Tabellen nedenfor viser målingerne af bremseevne, målt som antal millisekunder, det har taget at komme igennem svinget. Jo

færre millisekunder, desto bedre bremseevne.

1. Lav et parallelt boksplot (kassediagram) for hver faktor, der viser forskelle i bremseevnen for faktorens niveauer. Beskriv diagrammerne.

Da det er Magnussen, hvis bremseevne skulle forbedres, uden at Hülkenbergs bremseevne skulle forringes, samt at det ses ved et hurtigt skimt, at Hülkenbergs bremseevne ikke forringes, så tager jeg kun udgangspunkt i data fra Magnussen.

Et billede, der indeholder Rektangel, skærmbillede, kvadratisk, design

Automatisk genereret beskrivelseEt billede, der indeholder Rektangel, skærmbillede, design

Automatisk genereret beskrivelseEt billede, der indeholder Rektangel, diagram, kvadratisk, design

Automatisk genereret beskrivelseDa det er fabrikantens dæk jeg ønsker at undersøge, så laver jeg 3 boxplots, en til hver type dæk.

Figure 1: Hard

Figure 2: Medium

Figure 3: Soft

X Breaking systems ses øverst, brembo ses nederst.

Dataene er taget fra begge kørerer, så dermed 10 punkter til hver boxplot ( 5 magnussen, 5 Hülkenberg )

Mindst er bedst. For medium dæk så kan Brembo stadigvæk leverer det hurtigste sving, men nok kun for Hülkenberg, og ikke Magnussen.

================================================================================

Den generelle tendens er, at X Breaking systems leverer en mere consistent opbremsning blandt begge kørere. Der er altså noget mindre forskel på Magnussen og Hülkenberg med de nye bremserer.

================================================================================

### Lav en variansanalyse (ANOVA) på signifikansniveau 5 %, der belyser, om der mht. bremseevne er signifikant effekt af hver af de tre faktorer. Er der signifikante 2-faktor og 3-faktor interaktioner? Kom tilbage hertil.

1. Hvor mange frihedsgrader er der for testen for interaktion imellem de to faktorer bremsefabrikat og dæktype (A og C), og hvordan er antal frihedsgrader beregnet?
2. Lav en parvis sammenligning af bremseevnen for de tre dæktyper med Tukey HSD metoden. Hvilke dæktyper er forskellige på 5 % signifikansniveau? Hvad er 95 % konfidensintervallet for de parvise forskelle på bremseevne for de tre dæktyper?
3. Hvilke antagelser for residualerne er der gjort i denne statistiske model? Undersøg med plots, om antagelserne holder.

## Opgave 2. Bremsernes kølesystem

Bremseevnen er ikke den eneste parameter, der er vigtig i valget af bremsefabrikat. Bremsernes evne til at køle bremseskiverne ned igen efter en opbremsning er lige så vigtig. Bremserne virker ved at overføre racerbilens bevægelsesenergi til friktion (varme). Ved en hård opbremsning stiger bremsernes temperatur med ca. 100 °C pr. tiendedel sekund, og temperaturen kan overstige 1000 °C efter en opbremsning. Bremserne virker bedst mellem 400 og 800 °C, og de tager skade over 1000 °C. Racerbilens kølesystem til bremserne skal effektivt fjerne varmen fra bremseskiverne, så de ikke tager skade, og så bremserne er klar til næste nedbremsning.

Ingeniørerne fra MoneyGram Haas F1 Team har målt, hvor godt bremserne fra X Breaking Systems køler ned. For 800 nedbremsninger har de målt nedkølingen som temperaturfaldet ∆𝑡 i det første sekund efter bremsepedalen slippes. De 800 målinger af nedkølingen ∆𝑡 fordeler sig som vist i tabellen nedenfor.

Et billede, der indeholder tekst, Font/skrifttype, skærmbillede, nummer/tal

Automatisk genereret beskrivelse

Ingeniørerne antager, at bremsernes køleevne, målt som ∆𝑡, kan modelleres statistisk som en normalfordeling. De har estimeret normalfordelingens middelværdi og standardafvigelse ud fra de 800 målinger til at være henholdsvis 𝜇̂ = 229 °C og 𝜎􏰀 = 41 °C. De ønsker at benytte en Goodness of Fit test til at afgøre, om antagelsen om normalfordeling er plausibel.

1. Beregn – under forudsætning af ingeniørernes antagelse af en normalfordelt statistisk model – sandsynlighederne for at nedkølingen efter en opbremsning ligger indenfor temperaturintervallerne i tabellen. Du skal altså beregne de 5 sandsynligheder: P(∆t < 150 °C), P(150 °C ≤ ∆t < 200 °C), P(200 °C ≤ ∆t < 250 °C), P(250 °C ≤ ∆t < 300 °C), P(∆t ≥ 300 °C).

Så jeg kender den antagede middelværdi og den antagede spredning.

Så kan jeg finde procenterne i r ved at se på den kummerede normalfordeling fra

Når det er procenterne jeg leder efter, kan jeg bruge *r*’s *pnorm* funktion.

Jeg laver en funktion til at lave intervallerne. Som standard er pnorm sat til at være fra den nedre hale af intervallet. Til sidste interval sætter jeg den til at være den fra den øvre hale af intervallet og ned til 300.

Et billede, der indeholder tekst, skærmbillede, Font/skrifttype

Automatisk genereret beskrivelse

Og så har jeg sat det ind i en ny tabel:

Et billede, der indeholder tekst, skærmbillede, Font/skrifttype, nummer/tal

Automatisk genereret beskrivelse

1. Beregn under antagelse af ovennævnte normalfordeling det forventede antal nedbremsninger i hver af de fem temperaturintervaller, når der i alt foretages 800 nedbremsninger.

Nu har jeg procenterne, hvis jeg så siger

Hvor er det fulde antal på 800 nedbremsninger.

Det gør jeg for hver og sætter den i en ny tabel.

Et billede, der indeholder tekst, skærmbillede, Font/skrifttype, nummer/tal

Automatisk genereret beskrivelse

Tallene er ikke helt ca. tal, jeg rundede dem op for at slippe for decimaler.

1. Opstil formlen for teststørrelsen for en Goodness of Fit test og oplys teststørrelsens fordeling og antal frihedsgrader.

Her ønsker vi at bruge chi i anden fordelingen til som test af den fundne spredning.

Teststørrelsen vil være:

Hvor frihedsgraderne vil være

På baggrund af den estimerede spredning *s*

Og intervallet vil vi kunne finde ved:

1. Beregn den kritiske grænse for testen på 5 % signifikansniveau.

Spredningen finder jeg selv, da vi har et interval af målinger, hvor observationerne vægtes med antallet.

Jeg ganger antallet med middelværdien i intervallet, som jeg synes er en rimelig antagelse for et okay estimat.

Et billede, der indeholder tekst, Font/skrifttype, nummer/tal, skærmbillede

Automatisk genereret beskrivelse

For sidste interval fik jeg at vide, at man prøver at holde den under 1000° som er det kritiske punkt, så jeg tænker det er en okay antagelse.

Okay, jeg får ikke noget resultat ud af spredningen, jeg gør det ikke helt rigtigt.

Men lad mig regne chi i anden ud.

Uden spredningen finder jeg et estimat til:

=====================

=====================

1. Beregn teststørrelsens værdi og konkludér på testen.

## Opgave 3. Omgangstider

Ingeniørerne fra Haas F1 Team tester bremserne fra Brembo og X Breaking Systems (XBS) ved at måle omgangstider på en racerbane. Kevin Magnussen har kørt 20 omgange på en testbane med hvert af de to bremsefabrikater. I tabellen nedenfor vises omgangstiderne målt i sekunder med 3 decimalers nøjagtighed.

Et billede, der indeholder tekst, skærmbillede, nummer/tal, Font/skrifttype

Automatisk genereret beskrivelse

1. Beregn 95 % konfidensinterval for populationsmiddelværdien for omgangstider med hhv. Brembo og XBS bremserne. Hvad fortæller dine resultater dig om forskelle paÃä de to bremsefabrikater?

Her er jeg nødt til at estimerer middelværdien og spredningen. Den kommer så til at følge en t-fordeling med , frihedsgrader får både Brembo og XBS.

Konfidensintervallet beregnes derfra med middelværdien og dens estimeringsfejl.

Hvor

Og er standard afvigelsen, som vi kan beregne ud fra dataene.

Så nu har jeg alt, så er det bare om at beregne det.

Min kode:

Et billede, der indeholder tekst, skærmbillede, Font/skrifttype

Automatisk genereret beskrivelse

Resultatet:

Et billede, der indeholder tekst, Font/skrifttype, skærmbillede, hvid

Automatisk genereret beskrivelse



Så med sikkerhed kan jeg sige, at med Brembo ligger omgangstiderne mellem ca. 88 sekunder og 91,4 sekunder. Der er ikke meget variation her og det antages at være meget godt, når man skal køre mange omgange.

Med sikkerhed kan jeg sige, at med XBS ligger omgangstiderne mellem ca. 84,3 sekunder & 95,3 sekunder. Der er her meget variation, men mulighed for at køre meget hurtige omgange. Bremsesystemet vil måske kunne gøre det afgørende, hvis man kun skulle køre få omgange og var villig til at satse at tabe større, for muligvis at vinde stort.

1. Ingeniørerne fra Haas F1 Team vurderer, at der ikke er signifikant forskel paÃä de gennemsnitlige omgangstider for Brembo og XBS. Derimod vil de gerne foretage en hypotesetest for, om der på 5 % signifikansniveau er forskel paÃä varianserne. Opskriv nulhypotese og alternativhypotese for denne hypotesetest.

Så jeg ønsker at teste variansen. Til testen bruges F fordelingen, som bruges til test af varians for 2 stikprøver.

Jeg sætter nulhypotesen til at være ingen forskel i den egentlige spredning, og den alternative hypotese til at have en forskel mellem de to.

Hvis nulhypotesen er sand, så er forholdet mellem de to stikprøvers variation lige med 1.

Alternativ hypotesen kan vi bekræfte, hvis

Eller omvendt.

1. Opstil formlen for teststørrelsen og angiv hvilken fordeling, den følger.

Teststørrelsen følger F fordelingen, som tester varians af to stikprøver.

Ud fra nulhypotesen så antager jeg, at

Dermed simplificerer jeg

1. Beregn den kritiske region for testen, beregn teststørrelsens værdi og konkluder på hypotesetesten.

F fordelingen tager to frihedsgrader, en for hver stikprøve.

Begge stikprøver har 20 - 1 = 19 frihedsgrader.

er en funktion i *r* som tager procentsatsen, frihedsgraden på den ene, frihedsgraden på den anden, og beregner variansen ud fra det.

Med får jeg F værdien, og jeg skal bruge

Jeg får med 95% sikkerhed intervallet til at være:



For at varianserne skulle være ens, skulle forholdet mellem de to være på 1. 1 er langt udenfor intervallet og dermed har de to bremsere ikke samme varians på omgangstiderne.

========================================

Jeg kan dermed afkræfte nulhypotesen.

Dermed kan jeg så bekræfte alternativhypotesen,

de to bremsere varierer ikke lige meget.

========================================

1. Oplys hvilke antagelser, der er gjort i hypotesetesten, og om antagelserne er rimelige på baggrund af data.

Til hypotesetesten antog jeg, at deres rigtige spredning var ens. Dette var en rimelige antagelse, da det kun var, om der var forskel mellem de to, som vi ønskede at teste, og når antagelsen ikke holdte, så var der forskel mellem de to.