Norges teknisk-naturvitenskapelige universitet
Instittutt for matematiske fag

Anbefalte oppgaver 11, blokk II

Oppgavene i denne øvingen dreier seg om hypotesetesting og sentrale begreper som nullhypotese og alternativ hypotese, testobservator, forkastning og akseptanse, kritisk verdi, p-verdi, type 1 og type 2 feil, og teststyrke.

Oppgave 1

La $X_1, \ldots, X_5$ være uavhengige og normalfordelt med ukjent forventningsverdi μ og ukjent varians σ^2. Observerte verdier for $X_1, \ldots, X_5$ er:

i	1	2	3	4	5
x_i	5.00	4.48	5.60	4.25	5.44

a) Angi rimelige estimatorer for forventningsverdien μ og for variansen σ^2.

Hva blir estimatene med dataene gitt over ?

b) Utled et 95% konfidensintervall for μ basert på observasjonene over.

Forklar kort hvordan en fra dette konfidensintervallet kan avgjøre konklusjonen i en hypotesetest (signifikansnivå 5%) med alternativ hypotese $H_1 : \mu \neq \mu_0$ for ulike verdier for μ_0.

Oppgave 2

På ein av vegane inn til Trondheim er UP interessert i å måle effekten av ei holdingskampanje der målet var å få folk til å redusere farten på ei bestemt vegstrekning. På ein dag blei farten på 12 bilar målt. Vi skal gå utifrå at desse målingane er uavhengige og normalfordelte variable med forventing μ og standardavvik σ. Dei tolv observerte fartsmålingane er gitt nedanfor.

x_i: 75 61 85 65 69 82 70 67 62 93 77 74

Det oppgis at $\sum_{i=1}^{12} x_i = 880$ og $\sum_{i=1}^{12}(x_i - \overline{x})^2 = 1034.7$.

a) Forklar kva parameteren μ betyr i denne samanhengen. Skriv opp rimelege estimatorar for μ og σ i denne situasjonen. Kva blir estimata?

b) Forklar kva som meinast med *type 1 feil* når vi utfører ein hypotesetest.

Frå tidlegare undersøkingar har ein at gjennomsnittsfarten (av veldig mange bilar) var på 77 km/t. Tyder resultata frå desse målingane på at forventa fartsnivå på strekninga er lågare enn $77km/t$? Formuler spørsmålet som ein hypotesetest, gjennomfør testinga og gje konklusjonen. Bruk 5% signifikansnivå.

I punkt **c**) kan du gå utifrå at $\sigma = 10km/t$.

c) Forklar kva vi meiner med *type 2 feil* og kva som er samanhengen mellom denne og *styrken* til ein test. Gå utifrå at forventa fart til bilane er gått ned til $74km/t$. Finn sannsynet for at vi i testen i **b**) vil påstå at forventa fart til bilane er blitt lågare enn $77km/t$.

Finn deretter ut kor manga bilar vi må måle farten til for å få ein test som har styrke minst 0.90 når forventa fart $\mu = 74km/t$. Signifikansnivået skal framleis vere 5%.

Oppgave 3

Vi skal i denne oppgaven se for oss at vi kaster en mynt flere ganger. Mynten har 0.5 sannsynlighet for utfall «mynt» og 0.5 sannsynlighet for «kron». Vi antar at utfallene av ulike myntkast er uavhengige.

a) Anta at vi kaster mynten 5 ganger.

Hva er sannsynligheten for å få 5 «kron»?

Hva er sannsynligheten for å få 3 «kron»?

Hva er sannsynligheten for å få minst 4 «kron» på rad, det vil si en sekvens (*streak*) av bare «kron» utfall som minst er av lengde 4?

b) Vi kaster mynten 30 ganger. Fordelingen til lengste sekvens med «kron» er vanskelig å regne ut. I stedet kan vi få en datamaskin til å generere 30 stokastiske og uavhengige myntkast. Vi registrerer lengste sekvens av «kron». Prosedyren gjentas B ganger, og resultatet er representativt for fordelingen for lengste sekvens av «kron», se Figur 1.

Anslå sannsynligheten for å få en lengste sekvens på 5 eller 6.

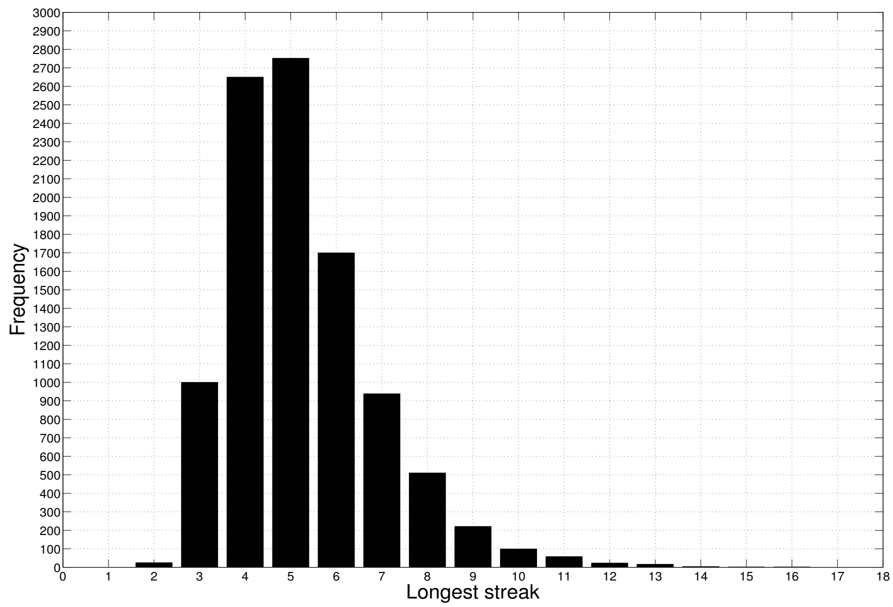
Miriam har fått hjemmelekse å kaste en mynt 30 ganger. Resultatet er som følger, der 0 betyr «mynt» og 1 betyr «kron»:
$(0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 1, 0, 0, 1)$.

Læreren mistenker at Miriam har jukset og bare funnet på tallene istedenfor å faktisk kaste en mynt, og læreren vil undersøke dette. Formuler dette som en hypotesetest om lengste sekvens av «kron». Bruk histogrammet i Figur 1 til å svare.

Oppgave 4

Produsenten av en bestemt bilmodell hevder at denne modellen kan forventes å kjøre minst 16 km pr. liter bensin på motorvei. Forbrukerorganisasjonen FO tester denne påstanden ved å kjøre et tilfeldig utvalg biler av denne modellen en passende distanse på en representativ motorvei og måle bensinforbruket.

På bakgrunn av erfaringer fra tidligere forsøk av samme type, antar FO at bensinforbruket til en tilfeldig valgt bil av den modellen som testes, kan modelleres med god tilnærmelse som en



Figur 1: Figuren viser et histogram av sekvenser. Disse er resultat av $B = 10000$ gjentak av 30 myntkast. Høyden på stolpene angir i hvor mange av de 10000 forsøkene lengste sekvens av ≪kron≫ var en bestemt lengde

normalfordelt tilfeldig variabel X med forventningsverdi μ og varians σ^2, dvs. $X \sim N(\mu, \sigma^2)$. Både forventningsverdien μ og standardavviket σ er i utgangspunktet ukjente størrelser.

Av praktiske grunner begrenser FO størrelsen på det tilfeldige utvalget til $n = 20$ biler. Etter forsøket ble alle målingene analysert, og resulterte i en gjennomsnittsverdi $\bar{x} = 15.56$ og et sample (empirisk) standardavvik $s = 0.94$.

a) Sett opp en hypotesetest for dette forsøket. La produsentens påstand representere nullhypotesen. Hvilken testobservator vil du bruke for å kontrollere hypotesen? Gi en kort begrunnelse for valget ditt. I forhold til et valgt signifikansnivå $\alpha = 0.05$, vil du akseptere produsentens påstand?

b) Finn P-verdien (signifikanssannsynligheten) for testen i punkt a) som svarer til de observerte verdiene.

Hvilken tilnærmelse kan du gjøre for at testobservatoren skal bli normalfordelt? Hvilken P-verdi får du hvis du bruker denne tilnærmelsen?

c) Bestem teststyrken for den alternative hypotesen $\mathrm{H}_1' : \mu = 15.5$ for signifikansnivå $\alpha = 0.05$ ved å bruke den samme normaltilnærmelsen som i punkt b). Gi et forslag til hvordan teststyrken kan økes.

Oppgave 5

Ved verdensmesterskap på enkeltdistanser på skøyter går hver deltager to 500 meters løp, ett løp hvor deltageren har indre bane i siste sving og ett løp hvor deltageren har ytre bane i siste sving. Rekkefølgen av deltagerne baserer seg på summen av tidene på de to løpene. Tilsvarende regel benyttes også i olympiske leker. Denne regelen ble innført fra og med verdensmesterskapet på Hamar i 1995. Tidligere ble rekkefølgen av deltagerne basert på kun et løp for hver deltager. Bakgrunnen for regelen om at hver deltager skal gå to løp er at det kan være en fordel å ha siste ytre siden løperne har stor fart i siste sving og i indre bane er krumningen større enn i ytre bane.

I et mesterskap med n deltagere, la Y_i og Z_i betegne løpstidene for løper nummer i for løpene med henholdsvis siste ytre og siste indre. La videre X være antall av de n løperne som har sin raskeste løpstid i løpet med siste ytre, dvs. X er antall løpere som har $Y_i < Z_i$. Vi antar at X er binomisk fordelt, dvs. $P(X = x) = b(x; n, p)$ der $p = P(Y_i < Z_i)$.

a) Angi hvilke forutsetninger som må være oppfylt i situasjonen beskrevet over for at antagelsen om at X er binomisk fordelt skal være korrekt.

Dersom $n = 20$ og $p = 0.7$, finn sannsynlighetene

$$P(X \leq 10) \quad \text{og} \quad P(X \geq 8 | X \leq 10).$$

b) Skriv opp rimelighetsfunksjonen (likelihoodfunksjonen) for p og benytt denne til å vise at sannsynlighetsmaksimeringsestimatoren for p blir

$$\widehat{p} = \frac{X}{n}.$$

Vis at $\widehat{p}$ er en forventningsrett estimator for p og at $\text{Var}(\widehat{p}) = p(1-p)/n$.

Videre i denne oppgaven skal vi benytte resultater fra 500 meter for menn i olympiske leker i Sochi i Russland i februar 2014 til å vurdere om det er grunnlag for å hevde at det er en fordel å gå siste ytre. Her var det $n = 39$ løpere som fullførte begge løpene, og av disse var det $x = 24$ som hadde sin raskeste løpstid i løpet med siste ytre.

I de videre utregningene kan du om nødvendig gjøre approksimasjoner, men du må i så fall begrunne disse.

c) Formuler en hypotesetest for situasjonen. Spesifiser H_0 og H_1, velg en passende testobservator og utled en beslutningsregel når signifikansnivået er $\alpha = 0.05$.

Hva blir konklusjonen av testen for resultatene fra Sochi.

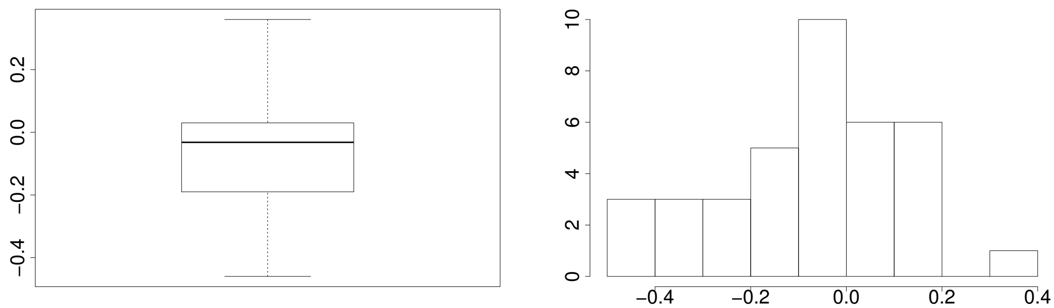
Regn også ut testens p-verdi basert på resultatene fra Sochi.

Vi skal i resten av denne oppgaven betegne testen du formulerte over som Test 1. Man kan formulere en alternativ test for samme situasjon, som vi skal kalle Test 2, ved å definere differansene

$$D_i = Y_i - Z_i \quad \text{for } i = 1, 2, \ldots, n,$$

der Y_i og Z_i altså er løpstiden for løper nummer i i løpene med henholdsvis siste ytre og siste indre. Man kan da lage en test ved å ta utgangspunkt i $\bar{D} = \frac{1}{n}\sum_{i=1}^{n} D_i$.

d) Formuler en hypotesetest for situasjonen med utgangspunkt i $\bar{D}$. Spesifiser H_0 og H_1, velg en passende testobservator og utled en beslutningsregel når signifikansnivået er α.



Figur 2: Boxplott og histogram over differansene d_i for de 37 løperne i 500 meter for menn i Sochi som ikke falt. I boxplottet er differansene langs y-aksen, mens i histogrammet er differansene langs x-aksen.

Av de 39 løperne i Sochi var det to som falt i et av sine løp. Dersom vi ser bort fra disse to løperne gir resultatene i Sochi $n = 37$, $\sum_{i=1}^{37} d_i = -2.654$ og $\sum_{i=1}^{37} d_i^2 = 1.552$. Hva blir da konklusjonen av Test 2 med resultatene fra Sochi når $\alpha = 0.05$? Avrund om nødvendig til nærmeste frihetsgrad oppgitt i tabell.

Det oppgis at teststyrken (sannsynligheten for å forkaste H_0 når H_1 er riktig) for Test 2 når $n = 37$, $E(D_i) = -0.07$ og $\text{Var}(D_i) = 0.2^2$ er lik 0.67. Disse verdiene for $E(D_i)$ og $\text{Var}(D_i)$ er cirka hva man får når man estimerer basert på resultatene av de 37 løperne i Sochi som ikke falt. Det kan også nevnes at med disse verdiene for $E(D_i)$ og $\text{Var}(D_i)$ blir $p = P(Y_i < Z_i) = 0.64$ hvis man antar at D_i er normalfordelt.

e) Finn teststyrken for Test 1 når $n = 39$ og $p = 0.64$.

Figur 2 viser boxplot og histogram over d_i for de 37 løperne i Sochi som ikke falt. Diskuter basert på konklusjonene du fant for Test 1 og Test 2, de utregnede teststyrkene for disse testene, samt plottene i Figur 2, hva du totalt sett ville konkludert med i den aktuelle situasjonen. De to løperne i Sochi som falt, falt begge i løpet hvor de hadde siste indre, har dette noen betydning for din konklusjon?

Fasit

1. **b**) $[4.23, 5.68]$

2. **a**) $\overline{x} = 73.3, s = 9.7$ **b**) $H_0 : \mu = 77$ mot $H_1 : \mu < 77$, Forkaster ikke H_0 **c**) Må måle farten på 96 bilar eller fleir

3. **a**) 0.031,0.313,0.09 **b**) 0.44, forkast H_0

4. **a**) Forkaster H_0 **b**) 0.025, 0.0183 **c**) 0.767

5. **a**) 0.048, 0.98 **c**) Ikke forkast H_0, 0.075 **d**) Forkast H_0 **e**) 0.545