

Vejledende eksamensopgaver vedr. hypotesetest (stx B og stx A)

Opgave 1

I nedenstående tabel ses resultaterne af samtlige hjerteklapoperationer i 2007-08 ved Odense Universitetshospital (OUH) sammenlignet med resultaterne ved landets øvrige hjertecentre.

	Død indenfor 30 dage efter operation	Overlevet de første 30 dage
OUH	11	206
De øvrige hjertecentre	32	1374

- a) Undersøg på 5%-signifikansniveau, om dødelighed indenfor 30 dage efter en hjerteklapoperation er ens ved OUH og de øvrige hjertecentre.
- b) Hvad ville resultatet være blevet, hvis vi havde valgt et signifikansniveau på 1%? Inddrag det statistiske begreb skjulte variable i din kommentar til resultatet.

Kilde: <http://www.si-folkesundhed.dk/upload/dhr2008.pdf>

Opgave 2

I en opinionsundersøgelse i X-købing er 500 personer blevet spurgt, om man er for eller imod de ændrede åbningstider i svømmehallen.

	For	Imod	Ved ikke	Sum
Kvinder	151	79	34	264
Mænd	115	70	51	236
Sum	266	149	85	500

Det oplyses, at man har en hypotese om, at de to køn har samme indstilling.

- a) Udfyld, under antagelse af at hypotesen er sand, nedenstående tabel over den forventede fordeling.

	For	Imod	Ved ikke	Sum
Kvinder				264
Mænd				236
Sum	266	149	85	500

- b) Bestem χ^2 -teststørrelsen, og den tilsvarende p-værdi, og kommenter resultatet.

Opgave 3

I et forsøg har nogle forskere undersøgt sammenhængen mellem indtagelse af pesticidet, Avadex, og forekomsten af svulster hos mus. I forsøget indgik i alt 95 mus, hvoraf 16 fik Avadex, mens de sidste 79 mus fungerede som kontrolgruppe. Resultatet blev, at 4 af de mus, der fik Avadex, udviklede svulster, og 5 af musene i kontrolgruppen udviklede svulster i den samme periode.

- a) Opstil en 2x2-tabel over forsøget, og angiv antallet af frihedsgrader.
- b) Undersøg på et 5% signifikansniveau om indtagelse af Avadex har betydning for udvikling af svulster hos mus.

Kilde: <http://www.matfys.kvl.dk/stat/kurser/statdata1/forel/week7-1-handout.pdf>

Opgave 4

I en klinisk afprøvning af en bestemt type medicin på en bestemt alvorlig sygdom, har man opdelt de lige behandlingskrævende patienter i to grupper A og B.

Gruppe A får medicinen, mens gruppe B er kontrolgruppe, som derfor får placebo (fx kalktabletter). Ingen af grupperne ved, hvilken type behandling de får.

Efter 6 måneders behandling undersøger man virkningen af behandlingen i hver af de to grupper, og resultatet af denne undersøgelse ses i tabellen nedenfor.

Status efter 6 mdr.	Gruppe A	Gruppe B	I alt
Død	8	17	25
Overlevende	75	48	123
I alt	83	65	148

I det følgende antages, at der ikke er nogen sammenhæng mellem medicineringen og patienternes overlevelseschancer.

- Opstil en tabel over de forventede værdier for "Død" og "Overlevende" i de to grupper A og B.
- Undersøg om medicineringen kan siges at have betydning for patienternes overlevelseschancer når signifikansniveauet fastlægges til 5%?

Opgave 5

Hjerte- og lungeoperationer kan medføre forstoppelse (postoperativ obstipation). På et hospital er man interesseret i at undersøge, om problemet har samme omfang ved de to operationstyper. De opstiller dette som en hypotese, de vil teste, og beder alle hjemsendte hjerte- og lungepatienter svare på følgende spørgsmål:

"Har du efter operationen haft forstoppelse i en grad, der har påvirket dine daglige gøremål?"

Patienternes svar fremgår af tabellen nedenfor, hvor tallene i parentes angiver de forventede værdier for patienternes svar.

Operationstype \ Problemer	Ja	Nej	Total
Hjerteoperation	9 (12,97)	51(47,03)	60
Lungeoperation	15(11,03)	36 (39,97)	51
Total	24	87	111

- Gør rede for, hvordan de forventede værdier er beregnet.
- Bestem χ^2 -teststørrelsen, og den tilsvarende p-værdi og kommenter resultatet.

Kilde: http://www.nhv.se/upload/dokument/forskning/Publikationer/MPH/2008/MPH2008-8_L%20S%20Rasmussen.pdf

Opgave 6

I en undersøgelse på en skole har en gruppe elever bedt 179 elever om at angive køn og rygevaner. En optælling viser, at der 36 rygere og 143 ikke-rygere blandt de adspurgte. Endvidere viser optællingen, at der ud af de 179 var 78 piger og 101 drenge. Det antages, at det er uafhængigt af køn, hvorvidt en elev ryger eller ej.

- Bestem under denne antagelse det forventede antal rygere blandt drenge, og udfyld de tomme pladser i skema 1 over den forventede fordeling af eleverne på rygevaner og køn:

Skema 1: Den forventede fordeling af elever på rygevaner og køn			
	Ryger	Ikke-ryger	
Hunkøn			78
Hankøn			101
Sum	36	143	179

I skema 2 ses stikprøvens faktiske fordeling af eleverne på køn og rygevaner:

Skema 2: Stikprøvens fordeling af elever på rygevaner og køn			
	Ryger	Ikke-ryger	
Hunkøn	21	57	78
Hankøn	15	86	101
Sum	36	143	179

b) Bestem χ^2 -teststørrelsen og den tilsvarende p-værdi

I en lignende undersøgelse kommer man frem til, at χ^2 -teststørrelsen er lig med 6,34.

c) Giver denne undersøgelse belæg for at hævde, at rygevaner er uafhængige af køn, når der benyttes et signifikansniveau på 5 %?

Opgave 7

En klasse vil undersøge, om drikkevaner er uafhængige af køn. Drikkevanerne inddeles i tre grupper: Driker ofte, drikke en gang i mellem, drikker aldrig. Resultatet af elevernes undersøgelser fremgår af følgende tabel:

335 elevers fordeling på drikkevaner				
	Driker ofte	Driker af og til	Driker ikke	
Pige	18	43	10	71
Dreng	48	175	41	264
	66	218	51	335

a) Giver resultatet belæg for at antage, at drikkevaner er uafhængige af køn, når der vælges et signifikansniveau på 5 %?

b) Undersøgelsen kritiseres for, at inddelingen mellem 'driker ofte' og 'driker af og til' er uklar. Optil selv ud fra tallene en ny tabel over de 335 elever hvor du kun anvender to inddelingskriterier, og undersøg om man når til samme konklusion nu som man gjorde i spørgsmål a.

Opgave 8

For en bestemt storproducent af legetøjsbolde gælder at diameteren af legetøjsboldene opfylder

Mindre end 20cm	Mellem 20cm og 22cm	Over 22 cm
10%	85%	5%

En forsendelse på 200 legetøjsbolde stoppes i tolden. Fordelingen af diameteren af disse legetøjsbolde er

Mindre end 20cm	Mellem 20cm og 22cm	Over 22 cm
28 stk	160 stk	12 stk

- a) Benyt statistiske metoder til at undersøge om forsendelsen kan stamme fra storproducenten

Opgave 9

I et eksperiment krydser en studerende lyserøde blomster med hinanden. Ifølge arvelighedslovene skulle det forventede resultat være 25 % røde, 50 % lyserøde og 25 % hvide. Resultatet af eksperimentet blev 236 blomster, der fordelte sig på farverne som vist i tabellen:

Eksperimentets fordeling af 236 blomster på tre farver			
Farve	rød	lyserød	hvid
hyppighed	66	115	55

- a) Beregn ud fra arvelighedsloven de forventede hyppigheder for de tre farver.
- b) Bestem χ^2 -teststørrelsen. Skriv et forslag til hvilken konklusion, den studerende kan skrive i sin rapport, når der anvendes et signifikansniveau på 5 %?

Opgave 10

I klageafdelingen i et firma modtager man klager pr. telefon. Firmaet har erfaring for, at klagebehandlingstiden pr. klage fordeler sig som angivet i nedenstående tabel.

Antal minutter	0-5	5-10	10-15	>15	I alt
Andel af det samlede antal klager	30%	40%	20%	10%	100%

Ved en kontroloptælling i en bestemt måned, viste det sig, at klagebehandlingstiden pr. klage fordelte sig som vist i tabellen nedenfor.

Klagebehandlingstid	0-5	5-10	10-15	>15	I alt
Observeret	37	53	25	5	120

- a) Hjælp firmaet med at opstille en nulhypotese, der kan testes. Undersøg dernæst om man på et 5% signifikansniveau kan sige, at mønstret i klagebehandlingstiden har ændret sig.

Kilde:

http://www.business.aau.dk/hd/Statistik/Dokumenter%20Rie/Hold%20A3_09/lektion%207/Kap%20H%20Goodness%20of%20fit%20test.doc

Opgave 11

Ved folketingsvalget i 2007 fordelte stemmerne sig således:

Parti	S	Rad	Kons	SF	DF	V	EL	Lib. All.	Kr. Dem.
Stemmer i %	25,5	5,1	10,4	13,0	13,8	26,3	2,2	2,8	0,9

Ved en meningsmåling med 968 respondenter foretaget af Greens Analyseinstitut 1/2 - 3/2 fordelte stemmerne sig således

Parti	S	Rad	Kons	SF	DF	V	EL	Lib. All.	Kr. Dem.
Stemmer i %	26,3	5,4	9,6	16,9	14,2	24,0	2,1	1,1	0,5

Vi vil anvende meningsmålingen til at undersøge følgende nulhypotese: stemmefordelingen er uændret in forhold til valget.

- a) Hvad der er populationen og hvad der er stikprøven. Lav en tabel over de observerede antal stemmer og de forventede antal stemmer under nulhypotesen.
- b) Bestem teststørrelsen, der tester om hypotesen skal accepteres eller forkastes, og afgør på et 5% signifikansniveau om man kan sige, at stemmefordelingen har ændret sig.
- c) Det ser ud som om, partiet SF har haft betydelig fremgang. Opstil en 2x2 tabel inddelt efter SF og øvrige partier med observerede og forventede værdier og afgør med statistiske metoder om SF's fremgang er signifikant på 5% niveau.