Stockholms universitet

OBS! Läs noga igenom anvisningarna i tentamen, t.ex. hur du ska skriva svaren. Det är ditt ansvar som student att följa de anvisningar som ges.
NOTE! Read the examination instructions carefully, e.g. how to write the answers. It is your responsibility as a student to follow the given instructions.

Skriv din anonymiseringskod och dagens datum på allt material du lämnar in.
(*Enter your anonymization code and todays date on all submitted materials*)

Anonymiseringskod (*Anonymization code*)	3	1	1	-	0	0	0	8	-	F	J	Y

Datum (*Date* YYYY-MM-DD)	2022-01-14	Plats nr. (*Seat No.*)	93

Kurs/Kurskod (*Course/Course code*)	ST745A
Kursmoment (*Course component*)	Statistisk inferens

Fylls i av tentamensvärd (*To be filled in by invigilator*)

Direkt i skrivning: (kryss)		Svarsblankett: (kryss)		Lösa svarsblad: (antal)	7
Lämnat in blankt: (kryss)		Dator: (kryss)			

Inlämningstid: 13 : 00 Signatur tentamensvärd:

Fylls i av lärare/examinator (*To be filled in by teacher/examinator*)

Betyg:	C	Poäng:	75

Signatur rättande lärare/examinator:

Stockholms universitet

Datum: 2022-01-14 (Date YYYY-MM-DD)
Kurs/Kurskod: ST745A (Course/Course code)
Sidnr.: (Page no.) 1
Anonymiseringskod (Anonymization code): 311-0008-FJY
Uppg.nr.: (Task no.) 1, a)/b)
Lärarens kommentar: (Teacher's note)

1. a) $f(x_1, x_2, \dots, x_n | \theta) = \theta^n \prod_1^n x_i^{\theta - 1}$ (independent random sample)

$= \theta^n e^{(\theta-1)\sum_1^n \ln x_i} = \theta^n \cdot e^{-\sum \ln x_i} e^{\theta \sum_1^n \ln x_i}$

$\Rightarrow$ exponential family with $c(\theta) = \theta^n$, $h(x) = e^{-\ln \sum x_i}$ ~~gh~~

$w(\theta) = \theta$, $t(x) = \sum_1^n \ln x_i$

And thus $\sum_1^n \ln x_i$ is a sufficient statistic for θ.

alternativt $e^{\sum \ln x} = \prod x_i$ som är en 1-1 transformation av din statistika

1. b)

$\theta a x^{a-1} e^{-\theta x^a}$, $x > 0$, $\theta > 0$, $a > 0$

joint distribution: $(\theta a)^n \prod_1^n x^{a-1} e^{-\theta \sum x_i^a}$ (independent random sample)

Using the factorization theorem

$f(x|\theta) = g(T(x)|\theta) h(x)$

We have that

$h(x) = a^n \prod x_i^{a-1}$

$g(T(x)|\theta) = \theta^n e^{-\theta \sum x_i^a}$

and our sufficient statistic is $T(x) = \sum_1^n x_i^a$. R

Poäng: (Points) 2

Tentamensservice/SU/typ_a

Uppg.nr.: (*Task no.*) 1 c)

Lärarens kommentar: (*Teacher's note*)

1. c) $f(x|\theta) = \dfrac{\theta a^\theta}{x^{\theta+1}}, \quad x > a, \ \theta > 0, \ a > 0$

$$f(x_1,\dots,x_n;\theta) = I_{(x>a)}\, \theta^n a^{n\theta} \prod x^{-(\theta+1)}$$

$$= I_{(x>a)}\, \theta^n a^{n\theta} \cdot e^{\ln \prod x_i^{-(\theta+1)}}$$

$$= I_{(x>a)}\, \theta^n a^{n\theta}\, e^{-(\theta+1)\sum_1^n \ln x_i}$$

$$= I_{(x>a)}\, e^{-\sum \ln x_i} \cdot \theta^n a^{n\theta}\, e^{-\theta \sum_1^n \ln x_i}$$

$\Rightarrow$ In exponential family form; ok

$$h(x) = I_{(x>a)}\, e^{-\sum \ln x_i}, \quad c(\theta) = \theta^n a^{n\theta}$$

$$w(\theta) = -\theta, \quad t(x) = \sum_1^n \ln x_i$$

And thus the sufficient statistic is

$$\sum_1^n \ln x_i$$

Alternativt se kommentar för uppg a.

Poäng: (*Points*)

Datum: (Date YYYY-MM-DD)	2022-01-14	Kurs/Kurskod: (Course/Course code)	ST745A	Sidnr.: (Page no.)
Anonymiseringskod (Anonymization code)	311-0008-FJY			2

Uppg.nr.: (Task no.) 2

2. For each of the colors, the experiment can be seen as N bernoulli trials.

For instance, the probability of x number of persons choosing their cinnamon bun from the red bowl is described by $P_R^{\Sigma x_i}(1-P_R)^{n-\Sigma x_i}$, $x = 0,1,2,\ldots,N$

$0 \leq P_R \leq 1$

Men du har en simultan fördelning här (multivariat!) (-5)

This is also the Likelihood-function.

To find the maximum-Likelihood estimator we differentiate the logL and set it equal to 0.

$$\ln L(P_R) = \Sigma x_i \ln P_R + (n-\Sigma x_i)\ln(1-P_R)$$

$$\frac{d\ln L}{dP_R} = \frac{\Sigma x_i}{\hat{P}_R} - \frac{(n-\Sigma x_i)}{1-\hat{P}_R} = 0$$

$$\Rightarrow \frac{\Sigma x_i}{n-\Sigma x_i} = \frac{\hat{p}}{1-\hat{p}} \Rightarrow \frac{(1-\hat{p})\Sigma x_i}{n-\Sigma x_i} - \hat{p} = 0$$

$$\Rightarrow \frac{\Sigma x_i}{n-\Sigma x_i} = \frac{\hat{p}\,\Sigma x_i}{n-\Sigma x_i} + \hat{p} \Rightarrow \hat{p}\left(1+\frac{\Sigma x_i}{n-\Sigma x_i}\right) = \frac{\Sigma x_i}{n-\Sigma x_i}$$

$$\Rightarrow \hat{P}_R = \frac{\frac{\Sigma x_i}{n-\Sigma x_i}}{1+\frac{\Sigma x_i}{n-\Sigma x_i}} = \frac{\frac{\Sigma x_i}{n-\Sigma x_i}}{\frac{n-\Sigma x_i+\Sigma x_i}{n-\Sigma x_i}} = \frac{\Sigma x_i(n-\Sigma x_i)}{n(n-\Sigma x_i)} = \frac{\Sigma x_i}{n}$$

$$\hat{P}_R = \bar{X}_R$$

where $n = N$ = number of persons in the experiment.

Poäng: (Points) 15

Uppg.nr.: (*Task no.*) 2

Lärarens kommentar: (*Teacher's note*)

2. a) Cont'd.

Now we check that we have found the max by taking the second derivative.

$$\frac{d\ln L}{dp_R} = \frac{\sum x_i}{p} - \frac{(n-\sum x_i)}{1-p}$$

$$\frac{d^2\ln L}{dp_R^2} = \frac{-\sum x_i}{p^2} - \frac{1-p+n-\sum x_i}{(1-p)^2}$$

$$= \frac{-\sum x_i}{p^2} - \frac{1}{1-p} - \frac{n-\sum x_i}{(1-p)^2}$$

Substituting p for it's ML-estimator $\frac{\sum x}{n}$:

$$\Rightarrow \frac{-n\bar{x}}{(\bar{x})^2} - \frac{1}{1-\bar{x}} - \frac{n-n\bar{x}}{(1-\bar{x})^2} = -\frac{n}{\bar{x}} - \frac{1+n}{1-\bar{x}} < 0$$

for $\bar{x} < 1$

$\Rightarrow \hat{p}_R = \bar{x}_R$ is the ML-estimator.

And thus, equivalently,

$\hat{p}_B = \bar{x}_B$ and $\hat{p}_G = \bar{x}_G$

2 b)

$$E(\hat{p}_R) = E\left(\frac{\sum x_i}{n}\right) = \text{independent} = \frac{1}{n}\sum E(x) = \frac{n}{n}\cdot E(x)$$

$= p_R$ (x is bernoulli distributed)

The same holds for $E[\hat{p}_B] = p_B$ and $E[\hat{p}_G] = p_G$

Poäng: (*Points*)

2 b) cont'd

$$Var(\hat{P}_R) = Var\left(\frac{\Sigma x_i}{n}\right) = \frac{1}{n^2} var(\Sigma x_i)$$

$$= \text{independent} = \frac{1}{n^2} \Sigma var(x) = \frac{1}{n} \cdot P_R(1-P_R)$$

$$= \frac{P_r(1-P_r)}{n}$$ (since X_R is bernoulli distributed)

correspondently,

$$Var(\hat{P}_B) = \frac{P_B(1-P_B)}{n}$$

ok

$$Var(\hat{P}_G) = \frac{P_G(1-P_G)}{n}$$

where $n = N$ = number of persons in the experiment.

Uppg.nr.:
(*Task no.*)

Lärarens kommentar:
(*Teacher's note*)

Poäng:
(*Points*)

Datum: (Date YYYY-MM-DD)	2022-01-14	Kurs/Kurskod: (Course/Course code)	ST745A	Sidnr.: (Page no.)
Anonymiseringskod (Anonymization code)	311-0008-FJY			4

Stockholms universitet

Uppg.nr.: (Task no.) 3

Lärarens kommentar: (Teacher's note)

3. a) $f(x;p) = (1-p)^{x-1} p \quad , \quad 0 < p < 1 \quad , \quad x = 1, 2, 3$

$$L(p;x) = (1-p)^{\sum x - n} p^n$$

$$\ln L = \left(\sum_1^n x_i - n\right) \ln(1-p) + n \ln p$$

$$\frac{d \ln L}{dp} = -\frac{(\sum x_i - n)}{1-\hat{p}} + \frac{n}{\hat{p}} = 0$$

$$\Rightarrow -\hat{p}(\sum x_i - n) + (1-\hat{p})n = 0$$

$$\Rightarrow -\hat{p}(\sum x_i - n) + n - \hat{p}n = 0$$

$$\Rightarrow -\hat{p}(\sum x_i - n + n) = -n$$

$$\Rightarrow \hat{p}(\sum x_i) = n$$

$$\Rightarrow \hat{p}_{ML} = \frac{n}{\sum_1^n x_i}$$

4

b)

Gaussian approximation gives that

- $E\left(\frac{n}{\sum x_i}\right) = E\left(\frac{1}{\bar{x}}\right) \approx \frac{1}{\mu}$ where $\mu = \frac{1}{p}$

further, from CLT we know that

$$\sqrt{n}(\bar{x} - E(x)) \xrightarrow{d} N(0, V(x))$$

and the delta method gives:

$$\sqrt{n}\left(\frac{1}{\bar{x}} - \frac{1}{\mu}\right) \xrightarrow{d} N\left(0, \underbrace{V(x) \cdot [g'(\mu)]^2}_{\text{Gauss-approx variance}}\right)$$

Poäng: (Points)

Tentamensservice/SU/typ_a

3b) cont

Uppg.nr.: (*Task no.*) 3

Lärarens kommentar: (*Teacher's note*)

We have!

$$g(\bar{x}) = \frac{1}{\bar{x}}$$

so $g'(\bar{x}) = -\frac{1}{\bar{x}^2}$ and $g'(\mu) = -\frac{1}{\mu^2}$

and $(g'(\mu))^2 = \frac{1}{\mu^4} = p^4$

$$\Rightarrow \sqrt{n}\left(\frac{1}{\bar{x}} - p\right) \xrightarrow{d} N(0, p^2(1-p))$$ R 8

Since $V(X) = \frac{1-p}{p^2}$ for the geometric distribution.

3 c) $I(p) = -E_p\left[\frac{d^2 \ln f}{d^2}\right]$

$$f(x;p) = (1-p)^{x-1} p$$

$$\ln f(x;p) = (x-1)\ln(1-p) + \ln p$$

$$d \ln f = -\frac{(x-1)}{1-p} + \frac{1}{p}$$

$$\frac{d^2 \ln f}{d^2} = \frac{-x-1}{(1-p)^2} - \frac{1}{p^2}$$

$$E\left[\frac{d^2 \ln f}{d^2}\right] = \frac{-(E(x)-1)}{(1-p)^2} - \frac{1}{p^2} =$$

$$-\frac{1}{p(1-p)^2} + \frac{1}{(1-p)^2} - \frac{1}{p^2}$$

Poäng: (*Points*)

Stockholms universitet

Datum: (Date YYYY-MM-DD) 2022-01-14	Kurs/Kurskod: (Course/Course code) ST745A		Sidnr.: (Page no.) 5
Anonymiseringskod (Anonymization code)	311-0008-FJY		

Uppg.nr.: (Task no.) 3 c)

Lärarens kommentar: (Teacher's note) 8

3. c) cont

$$E\left[\frac{\partial^2 \ln f}{\partial p}\right] = -\frac{p}{p^2(1-p)^2} + \frac{p^2-(1-p)^2}{p^2(1-p)^2}$$

$$= \frac{p^2 - p - (1^2 - 2p + p^2)}{p^2(1-p)^2} = \frac{p-1}{p^2(1-p)^2} = \frac{1-p}{-p^2(1-p)^2}$$

$$= -\frac{1}{p^2(1-p)}$$

$$\Rightarrow I_{(p)} = \frac{1}{p^2(1-p)}$$ on

And $\frac{1}{I}$ is thus the ~~attained~~ asymptotically attained variance we already calculated through the delta method, the [illegible]

Poäng: (Points) 20

Tentamensservice/SU/typ_a

Uppg.nr.:
(*Task no.*)

Lärarens kommentar:
(*Teacher's note*)

Poäng:
(*Points*)

Datum: (Date YYYY-MM-DD) 2022-01-14
Kurs/Kurskod: (Course/Course code) ST745A
Sidnr.: (Page no.) 6
Anonymiseringskod (Anonymization code) 311-0008-FJY
Uppg.nr.: (Task no.) 4.
Lärarens kommentar: (Teacher's note)

4. a) I will to this by setting upp two test: $H_0: P_R = P_B = P_G$

$\Rightarrow H_{0_1}: P_R = \frac{1}{3}$, $H_1: P_R \neq \frac{1}{3}$

$H_{0_2}: P_B = \frac{1}{3}$, $P_B \neq \frac{1}{3}$

$$\lambda_1 : \frac{\sup L(H_{0_1})}{\sup L(H_1)} = \frac{\left(\frac{1}{3}\right)^{\sum x_i} \left(\frac{2}{3}\right)^{n-\sum x_i}}{\bar{x}^{\sum x_i}(1-\bar{x})^{n-\sum x_i}} < C$$

and corresponding for H_{0_2}.

b) $-2\ln\lambda$ asymptotically, as $n \to \infty$, $\to \chi^2_1$

So the tests becomes, Reject H_{0_1} if:

$$-2\ln\left(\frac{\frac{1}{3}^{\sum x_i} \frac{2}{3}^{n-\sum x_i}}{\bar{x}^{\sum x_i}(1-\bar{x})^{n-\sum x_i}}\right) > \chi^2_{(1)}\text{-critical}$$

value from table for chosen α

If rejected, reject H_0. otherwise also test $H_{0_2}: P_B = \frac{1}{3}$, $H_{1_2}: P_B \neq \frac{1}{3}$.

Poäng: (Points) 0

Uppg.nr.:
(*Task no.*)

Lärarens kommentar:
(*Teacher's note*)

Poäng:
(*Points*)

Stockholms universitet

Datum: (Date YYYY-MM-DD)	2022-01-14	Kurs/Kurskod: (Course/Course code) ST 745 A	Sidnr.: (Page no.) 7
Anonymiseringskod (Anonymization code)	311-0008-FJY		

Uppg.nr.: (Task no.) 5

Lärarens kommentar: (Teacher's note)

5. $X_1, \ldots X_n$ IID $N(\mu, \sigma^2)$ μ known

$H_0: \sigma \geq \sigma_0$

$H_A: \sigma < \sigma_0$

Neyman - Pearsons lemma says that the strongest test ~~for~~ for simple hypothesis $H_0: \theta = \theta_0$ and $H_1: \theta = \theta_1$ is to reject H_0 if

$L(\theta_0)/L(\theta_1) < C$.

In our case we would have

$$\frac{L(\sigma_0)}{L(\sigma_1)} = \frac{\left(\frac{1}{\sqrt{2\pi}\sigma_0}\right)^n e^{-\frac{\Sigma(x_i-\mu)^2}{2\sigma_0^2}}}{\left(\frac{1}{\sqrt{2\pi}\sigma_1}\right)^n e^{-\frac{\Sigma(x_i-\mu)^2}{2\sigma_1^2}}} < C$$

$$= e^{\Sigma(x_i-\mu)^2\left(\frac{1}{2\sigma_1^2} - \frac{1}{2\sigma_0^2}\right)} < \left(\frac{\sigma_0}{\sigma_1}\right)^n \cdot C$$

$$= \Sigma(x_i-\mu)^2\left(\frac{1}{2\sigma_1^2} - \frac{1}{2\sigma_0^2}\right) < n\ln\left(\frac{\sigma_0}{\sigma_1}\right) + \ln C$$

$$= \text{since } \sigma_1 < \sigma_0 = \Sigma(x_i-\mu)^2 < \frac{n\ln\frac{\sigma_0}{\sigma_1} + \ln C}{\frac{1}{2\sigma_1^2} - \frac{1}{2\sigma_0^2}}$$

But this holds for all $\sigma_1 < \sigma_0$

and $\Sigma(x_i-\mu)^2 < C'$ is an UMP

Poäng: (Points) 20

Uppg.nr.: (*Task no.*) 5

Lärarens kommentar: (*Teacher's note*)

5 b)

$$\sum (x_i - \mu)^2 < c'$$

$$=$$

Under H_0 $\sum \left(\frac{x_i - \mu}{\sigma_0} \right)^2$ is $\chi^2_{(n)}$-distributed

and thus Reject H_0 if $\sum \left(\frac{x_i - \mu}{\sigma_0} \right)^2$

is lower than the critical value in the

left tail of the $\chi^2_{(n)}$-distribution, where

! $P(X < c'') = \alpha$

Bra!

Poäng: (*Points*)