Rettevejledning økonometri A

Målbeskrivelse:

Kurset har som mål at introducere studerende til sandsynlighedsteori og statistik. Målet er, at de studerende efter at have gennemført faget kan:

- Forstå og benytte de vigtigste sandsynlighedsteoretiske begreber som: sandsynlighed, simultane-, marginale- og betingede- sandsynligheder, fordeling, tæthedsfunktion, uafhængighed, middelværdi, varians og kovarians samt at selvstændigt kunne anvende disse begreber på konkrete problemstillinger
- Kende resultatet fra den centrale grænseværdi sætning
- Kende og genkende de mest anvendte diskrete og kontinuerte fordelinger som: Bernoulli, Binomial, Poisson, multinomial, negative binomial fordeling, hypergeometrisk, geometrisk, lige-, normal-, Chi-i-anden-, eksponential, gamma-, t-, F-fordeling samt at selvstændigt kunne arbejde med disse fordelinger i konkrete problemstillinger
- Forstå de vigtigste statistiske begreber som: tilfældige udvælgelse, likelihood funtionen, sufficiens, stikprøvefunktion, egenskaber ved stikprøvefunktionen, estimation heruden maksimum likelihood og moment estimation, konsistens, konfidensinterval, hypoteseprøvning, teststørrelser, hypoteser, testsandsynlighed, signifikansniveau og type I og II fejl
- Være i stand til selvstændigt at gennemføre en simpel statistisk analyse, som involverer estimation, inferens og hypoteseprøvning, f.eks. sammenligning af middelværdien i to populationer eller uafhængighedstest for diskrete stokastiske variable.
- Indlæse og kombinere datasæt, lave nye variable, udtrække en stikprøve og udføre simple statistiske analyser ved hjælp af statistik-pakken SAS
- Beskrive resultatet af egne analyser og overvejelser i et klart og tydeligt sprog

Opgave 1

1. N=80.000, M = 1000 n = 1000. k er antallet af mærkede pengesedler. Denne følger en hypergeometrisk fordeling. P(k=j)=

2. $E[k] = n \cdot \frac{M}{N} = 13$ det betyder at $N = \frac{1000^2}{13} = 76923$

3. $P(k = 13|N = 76923, M = 1000, n = 1000) = 0,11138761185$, $P(k = 13|N = 76922, M = 1000, n = 1000) = 0,11138761183$, $P(k = 13|N = 76924, M = 1000, n = 1000) = 0,11138761162$. Den fundne værdi af N maksimerer sandsynligheden for at finde 13 mærkede sedler.

Opgave 2

1. $E[Y] = 25,83$ og $Var(Y) = 23,78$. $Cov(X,Y) = 5,39$ og $\rho_{XY} = 0,41$. Dvs. jo længere uddannelse des højere løn.

2. Varians og kovarians beregnes lettest fra $\text{Y}^{årlig} = 12Y.Cov(12Y,X) = 12Cov(Y,X)$ og $\rho_{12YX} = \frac{12Cov(X,Y)}{\sqrt{Var(x)144Var(Y)}} = \rho_{XY}$. En skalering af en stok. variabel har indflydelse på kovariansen, men ikke på korrelationskoefficienten.

2. Skatten skal beregnes på årsniveau.

$\text{Y}^{årlig}$	T	$\text{Y}^{efterskat}$
264	85,6	178,4
336	121,6	214,4
420	172	248

Nu kan kovarians og korrelation mellem $\text{Y}^{efterskat}$ og X beregnes. $\text{Cov}(\text{Y}^{efterskat}, X) = 29,93$, $\rho_{Y^{efterskat}} = 0,23$. Efter skat indkomsten er ikke en lineær funktion (den er kun stykkevis lineær) af den stokastiske variabel Y fåes at korrelationskoefficienten er anderledes. (det kan også bemærkes at: Der er progressiv skat og det betyder at efterskat indkomsten er mindre korreleret med uddannelse end før skat indkomsten).

Opgave 3

1. Uafh. og samme p. parametre er n og p.

2. $\widehat{p} = \frac{300}{950}$, middelret og konsistent estimator

3. $\widehat{p} \in [\frac{300}{950} - 1,645\sqrt{\frac{\frac{300}{950}(1-\frac{300}{950})}{950}}, \frac{300}{950} + 1,645\sqrt{\frac{\frac{300}{950}(1-\frac{300}{950})}{950}}] = [0,291, 0,341]$, p er en andel og derfor bruges altid konfidensinterval for 'large samples'.

4. $H_0 : p = 0,25, H_a : p > 0,25, Z = \frac{\widehat{p} - 0,25}{\sqrt{\frac{0,25(1-0,25)}{950}}}, p = 1 - \Phi(Z) = 1,414e-6.$ H_0 kan afvises på 1 pct. niveau.

5. $H_0 : \widehat{p}^m = \widehat{p}^k, H_a : \widehat{p}^m \neq \widehat{p}^k, Z = \frac{\widehat{p}^m - \widehat{p}^k}{\sqrt{\widehat{p}(1-\widehat{p})(\frac{1}{n_1} + \frac{1}{n_2})}}, p = 2\Phi(Z) = 0,170.H_0$ forkastes ikke.

6. $H_0 : p_{j|i} = p_j, \forall j, i = m, k, H_a$: det er ikke tilfældet. $Z = 12,59\tilde{}\chi^2(2)$. p<0.001. Dvs. H_0 kan forkastes