Aalto-universitetet

MS-A0509 Grundkurs i sannolikhetskalkyl och statistik — Gripenberg
Övning 4
Vecka 5, 26–30.1.2015

Besvara Stack-uppgifterna (`stack3.aalto.fi/course/view.php?id=18`)
senast 2.2.2015 kl. 12.00.

Lämna in lösningarna till I-uppgifterna senast 2.2.2015 kl. 12.00.
Kom ihåg att skriva ditt namn och studentnummer!

I1. Vattenhöjden i Saramojoki var under åren 2008–2010 följande:

76	54	52	98	170	59	42	80	69
99	131	86	58	48	42	53	137	50
45	35	43	65	64	73	46	39	38
108	129	111	45	24	37	51	70	43

Bilda en frekvenstabell med klasserna $21-30$, $31-40$, $41-50$, $51-60$, $61-70$, $71-80$, $81-90$, $91-100$, $101-110$, $111-130$, $131-150$ och $151-170$ och rita ett histogram över materialet (och kom ihåg att areorna av rektanglarna skall vara proportionella mot frekvenserna).

Ledning: I storleksordning är siffrorna 24, 35, 37, 38, 39, 42, 42, 43, 43, 45, 45, 46, 48, 50, 51, 52, 53, 54, 58, 59, 64, 65, 69, 70, 73, 76, 80, 86, 98, 99, 108, 111, 129, 131, 137, 170.

I2. Använd samma data som i uppgift I1. Beräkna medianen, variationsbredden (största minus minsta), och avståndet mellan kvantilerna $x_{0.75}$ och $x_{0.25}$. Vilken av siffrorna 33.533, 68.611 och 1124.5 är medelvärdet, vilken varians och vilken standardavvikelse. (Du behöver inte siffrorna i uppgift I1 för att svara på den senare frågan.)

Om du skulle vara tvungen att beskriva siffermaterialet med två tal, vilka skulle du välja? Blir det en bra beskrivning med de två tal du valt? (Det finns inte ett entydigt rätt svar på dehär frågorna!)

I3. Antag att du har ett observerat stickprov x_j, $j = 1, \dots, 8$ av en slumpvariabel som du antar har fördelningen $\mathrm{N}(\mu, \sigma^2)$. Du har räknat ut medelvärdet $\overline{x}$ och stickprovsvariansen $s^2 = 2.6753$, och sedan på vanligt sätt bestämt ett (symmetriskt) konfidensintervall som blev $[2.6543, 6.7017]$. Vad är konfidensgraden?

I4. Antag att $X_i, i = 1, 2, \dots, n$ är oberoende slumpvariabler så att $\mathrm{E}(X_i) = 0$ och $\mathrm{Var}(X_i) = \sigma^2$ då $i = 1, 2, \dots, n$.
Visa att $\mathrm{E}(S^2) = \sigma^2$ då

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2,$$

och $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ dvs. stickprovsvariansen är en **väntevärdesriktig** estimator av variansene.

Ledning: Räkna $\mathrm{E}\big((n-1)S^2\big)$ *och observera att eftersom* $\mathrm{E}(X_i) = 0$ *så är* $\mathrm{E}(X_i^2) = \mathrm{Var}(X_i)$ *och* $\mathrm{E}(\overline{X}) = 0$ *så att* $\mathrm{E}(\overline{X}^2) = \mathrm{Var}(\overline{X})$ *och kom ihåg vad* $\mathrm{Var}(\overline{X})$ *blir då slumpvariablerna* X_i *är oberoende.*

Obs! *Om* $\mathrm{E}(X_i) = \mu \neq 0$ *så kunde vi byta ut slumpvariablerna* X_i *mot slumpvariablerna* $Y_i = X_i - \mu$ *utan att* S^2 *skulle ändras och detta betyder att antagandet att* $\mathrm{E}(X_i) = 0$ *inte är en begränsning, endast en förenkling.*

I5. Slumpvariablen X har täthetsfunktionen

$$f(t, \theta) = \begin{cases} 2^{\theta-1}(\theta - 1)t^{-\theta}, & t \geq 2, \\ 0, & t < 2, \end{cases}$$

där $\theta > 1$. Ett stickprov gav värdena 4, 6, 8 och 14. Bestäm ett estimat för θ med momentmetoden och ett med "maximum likelihood"-metoden.

Ledning: Kom ihåg att $\int_2^\infty t 2^{\theta-1}(\theta - 1)t^{-\theta}\,\mathrm{d}t = 2\frac{\theta-1}{\theta-2}$ *då* $\theta > 2$ *och att* $\frac{\mathrm{d}}{\mathrm{d}\theta}a^\theta = a^\theta \ln(a)$.

Svar: 2.3333 och 1.7806