Eksamen på Økonomistudiet vinteren 2017-18
Sandsynlighedsteori og Statistik
2. årsprøve
6. januar, 2018

(3-timers prøve med hjælpemidler)

Rettevejledning

Opgaven består af tre delopgaver, som alle skal besvares. De tre opgaver kan regnes uafhængigt af hinanden. Opgave 1 og 2 indgår tilsammen med samme vægt som opgave 3.

Opgave 1

I denne opgave undersøges det forventede afkast på en "call option". Antag at vi har et aktiv, som i dag har en pris på 100, men hvor der er usikkerhed om prisen på aktivet om et år. Vi antager, at prisen på aktivet om et år kan beskrives ved en kontinuert stokastisk variabel X. Vi antager, at X er ligefordelt på intervallet $[90; 120]$. Vi definerer indtjeningen som $Z = X - 100$.

1. Sandsynligheden for at prisen falder, og køberen herved taber penge, $P(Z < 0)$ kan beregnes ved at opskrive tæthedsfunktionen for Z. Da X er ligefordelt på intervallet $[90, 120]$, så følger, at Z er ligefordelt på $[-10, 20]$. Tæthedsfunktionen kan opskrives som

   $$f(z) = \frac{1}{20-(-10)}\mathbf{1}_{\{[-10,20]\}}(z) = \frac{1}{30}\mathbf{1}_{\{[-10,20]\}}(z)$$

   Sandsynligheden kan udregnes som

   $$P(Z < 0) = \int_{-\infty}^{0} \frac{1}{30}\mathbf{1}_{\{[-10,20]\}}(z)dz = \frac{1}{30}\int_{-10}^{0} dz = \frac{10}{30} = 1/3$$

   Opgaven kan også besvares ved grafisk at resonnere sig frem til sandsynligheden.

2. Den forventede indtjening $E(Z)$ kan udregnes ved at bruge middelværdiformlen for en ligefordeling

   $$E(Z) = \frac{b+a}{2} = \frac{20+-10}{2} = \frac{10}{2} = 5.$$

3. Den forventede indtjening, givet at indtjeningen er positiv $E(Z|Z > 0)$,

kan udregnes ved at bruge formlen for betingede middelværdier

$$\begin{aligned} E(Z|Z>0) &= \frac{\int\limits_{(z>0)} z\frac{1}{30}\mathbf{1}_{\{[-10,20]\}}(z)dz}{P(Z>0)} \\ &= \frac{\int\limits_{0}^{20} z\frac{1}{30}dz}{1-P(Z<0)} \\ &= \frac{\frac{1}{30}\left[\frac{1}{2}z^2\right]_0^{20}}{1-\frac{1}{3}} \\ &= \frac{\frac{1}{30}\frac{1}{2}20^2}{\frac{2}{3}} \\ &= \frac{3}{2}\frac{1}{30}\frac{1}{2}400 = 10 \end{aligned}$$

Opgaven kan også besvares ved grafisk at tegne den betingede sandsynlighedstæthed og herved indse, at det er en ligefordeling på intervallet [0,20].

Det er nu muligt at købe en "call option", som sikrer, at man har ret til at købe aktivet til den faste pris 100 (men ikke er forpligtet). Med denne "call option" kan man sikre sig mod negativ indtjening, idet man kun vil købe, hvis prisen overstiger 100. Indtjeningen, når man har en call option, er givet ved

$$Y = \max(0, Z).$$

4. Den forventede indtjening for call optionen $E(Y)$ kan beregnes ved at udnytte eksempel 12 i Anders Rahbeks note.

$$\begin{aligned} E(Y) &= E(Y|Y>0)\cdot P(Y>0) + E(Y|Y\leq 0)P(Y\leq 0) \\ &= E(Z|Z>0)\cdot P(Z>0) + E(0|Z<0)P(Z<0) \\ &= E(Z|Z>0)\cdot P(Z>0) + 0\cdot P(Z<0) \\ &= E(Z|Z>0)\cdot P(Z>0) \\ &= 10\cdot\frac{2}{3} = 6\frac{2}{3} \end{aligned}$$

Hvis prisen er 3, er det ikke en god ide at købe call optionen da

$$E(Y) - E(Z) = 6\frac{2}{3} - 5 = 1\frac{2}{3}$$

altså mindre end 3.

Opgave 2

Lad X være en stokastisk variabel der angiver om økonomien er i en krise ($X = 1$) eller ej ($X = 0$), med $P(X = 0) = 2/3$.

1. Find $E(X)$.

   Facit: $E(X) = 1 \cdot P(X = 1) = 1/3$.

2. Under en krise gælder der at sandsynligheden for stort tab på investeringer er større end hvis der ikke er krise. Det kan formaliseres ved at indføre endnu en stokastisk variabel Y med $Y = 1$ svarende til stort tab, og $Y = 0$ svarende til lille tab. De to variable X og Y er ikke uafhængige, men vi kan angive:

   $$\begin{aligned} P(Y = 1 | X = 1) &= 2/3 \\ P(Y = 1 | X = 0) &= 1/3. \end{aligned}$$

   Find $E(Y)$.

   Facit:

   $$\begin{aligned} E(Y) &= P(Y = 1) \\ &= P(Y = 1 | X = 1) P(X = 1) + P(Y = 1 | X = 0) P(X = 0) \\ &= \frac{2}{3}\frac{1}{3} + \frac{1}{3}\frac{2}{3} \\ &= \frac{4}{9}. \end{aligned}$$

3. Find $E(Y | X = 1)$.

   Facit: $E(Y | X = 1) = P(Y = 1 | X = 1) = 2/3$.

4. Find $E(XY)$.

Facit:

$$\begin{aligned} E(XY) &= E(XY|X=1)\,P(X=1) \\ &= E(Y|X=1)\,P(X=1) \\ &= \frac{2}{3}\frac{1}{3} = \frac{2}{9}. \end{aligned}$$

Opgave 3

1. Hvis man differentierer fordelingsfunktionen får man Weibull fordelingens tæthed:

   $$\frac{\partial}{\partial y} F_{Y_i}(y \mid \lambda) = \frac{\partial}{\partial y}\left(1 - \exp(-\lambda y^2)\right) = 2\lambda y \exp(-\lambda y^2) = f_{Y_i}(y \mid \theta).$$

   For at $F_{Y_i}(y \mid \lambda)$ er en fordelingsfunktion skal der desuden gælde, at

   $$F_{Y_i}(0 \mid \lambda) = 1 - \exp(-\lambda 0^2) = 0,$$

   og at, når $y \to \infty$, gælder der at

   $$F_{Y_i}(y \mid \lambda) = 1 - \exp(-\lambda y^2) \to 1.$$

2. Med den valgte fordelingsantagelse er likelihood bidraget givet ved

   $$\ell(\lambda \mid Y_i) = 2\lambda Y_i \exp(-\lambda Y_i^2),$$

   som er antaget identisk for alle $i = 1, 2, ..., n$. Så er log-likelihood bidraget givet som

   $$\log \ell(\lambda \mid Y_i) = \log(2) + \log(\lambda) + \log(Y_i) - \lambda Y_i^2.$$

   Under antagelse af uafhængighed er den samlede likelihood funktion givet som

   $$L_n(\lambda) = \prod_{i=1}^{n} \ell(\lambda \mid Y_i) = \prod_{i=1}^{n} 2\lambda Y_i \exp(-\lambda Y_i^2),$$

   sådan at

   $$\log L_n(\lambda) = \sum_{i=1}^{n} \left\{\log(2) + \log(\lambda) + \log(Y_i) - \lambda Y_i^2\right\}.$$

3. Score-bidraget er givet som

$$s_i(\lambda) = \frac{\partial \log \ell\left(\lambda \mid Y_i\right)}{\partial \lambda} = \frac{1}{\lambda} - Y_i^2,$$

så scoren er

$$S(\lambda) = \sum_{i=1}^{n} s_i(\lambda) = \sum_{i=1}^{n} \left\{ \frac{1}{\lambda} - Y_i^2 \right\}.$$

Dermed er første ordens betingelsen for maximum likelihood estimatoren givet ved

$$S(\hat{\lambda}_n) = \sum_{i=1}^{n} \left\{ \frac{1}{\hat{\lambda}_n} - Y_i^2 \right\} = \frac{n}{\hat{\lambda}_n} - \sum_{i=1}^{n} Y_i^2 = 0,$$

som løses hvor

$$\hat{\lambda}(Y_1, ..., Y_n) = \frac{n}{\sum_{i=1}^{n} Y_i^2}.$$

4. I det konkrete tilfælde med $n = 100$ og $\sum_{i=1}^{100} y_i^2 = 6153.4$ er estimatet givet ved

$$\hat{\lambda}(y_1, ..., y_n) = \frac{100}{6153.4} = 0.01625.$$

5. Hesse-bidraget fra Y_i er givet som

$$H_i(\lambda) = \frac{\partial^2 \log \ell\left(\lambda \mid Y_i\right)}{\partial \lambda \partial \lambda} = \frac{\partial s_i\left(\lambda\right)}{\partial \lambda} = \frac{\partial}{\partial \lambda}\left(\frac{1}{\lambda} - Y_i^2\right) = \frac{-1}{\lambda^2},$$

så informationen, evalueret i den sande værdi λ_0, er givet som

$$I(\lambda_0) = E(-H_i(\lambda_0)) = \lambda_0^{-2}.$$

Variansen på estimatoren er derfor

$$V(\hat{\lambda}_n) = \frac{1}{n} I(\lambda_0)^{-1} = \frac{\lambda_0^2}{100}.$$

Med estimatet indsat i stedet for λ_0 fås

$$V(\hat{\lambda}_n) \approx \frac{\hat{\lambda}_n^2}{100} = \frac{0.01625^2}{100},$$

og derfor er

$$\mathsf{se}(\hat{\lambda}_n) = \sqrt{\frac{0.01625^2}{100}} = 0.001625.$$

6. I opgaven er *Time-to-failure* for et *normalt* produkt defineret som medianen. Baseret på estimated bliver det

$$\sqrt{\log(2)/\hat{\lambda}_n} = \sqrt{\log(2)/0.01625} = 6.53.$$

7. En median på 8 måneder svarer til at $\sqrt{\log(2)/\lambda} = 8$, sådan at

$$\lambda = \log(2)/64 = 0.01083.$$

For at teste dette, anvendes hypoteserne

$$H_0 : \lambda_0 = 0.01083 \quad \text{mod} \quad H_A : \lambda_0 \neq 0.01083.$$

Teststørrelsen er givet som

$$z_n(\lambda_0 = 0.01083) = \frac{\hat{\lambda}_n - 0.01083}{\mathsf{se}(\hat{\lambda}_n)} = \frac{0.01625 - 0.01083}{0.001625} = 3.335.$$

Hvis nul-hypotesen er korrekt, er størrelsen fordelt som en standard normal-fordeling, $N(0,1)$, så på $\alpha = 0.05$ signifikans-niveau er den kritiske værdi 1.96. Hypotesen afvises klart. Man kan udregne p-værdien for dette tilfælde som

$$p_n(\lambda_0 = 0.01083) = 2\,(1 - \Phi(3.335)) = 0.000853.$$

8. Vi udregner direkte

$$\begin{aligned} P(Y_i > 10) &= 1 - P(Y_i \leq 10) \\ &= 1 - F_{Y_i}(10 \mid \hat{\lambda}_n = 0.01625) \\ &= 1 - (1 - \exp(-0.01625 \cdot 10^2)) \\ &= \exp(-0.01625 \cdot 10^2) \\ &= 0.197, \end{aligned}$$

svarende til 19.7%.