Institut for Matematik
AARHUS UNIVERSITY

Introduktion til Sandsynlighedsteori og Statistik
Markus Kiderlen
October 10, 2023

UGESEDDEL 8

Forelæsningerne i Uge 6: Vi har gennemgået Afsnit 4.1 og 4.2, samt teksten *Integrationsnoter*, der findes på Brightspace → Content → Supplerende Noter → Matematiske redskaber. Afsnit 4.3 er ikke pensum. Vi har løst *The Coupon Collector's Problem* (Brightspace→Content→Supplerende Noter) og jeg beskrev hvordan resultatet kan tolkes for *hash tabeller* i datalogi. Dermed har vi afsluttet Kap. 4.

Forelæsningerne i Uge 7:

- Forelæsning 13 (10. oktober): Vi starter med Kapitel 5. Jeg regner med at nå til Afsnit 5.1.4. Afsnit 5.1.5 er ikke pensum, pånær Lemma 5.2 og punkt 4. i boksen på side 242. Hvis der er tid vil vi starte med noterne *Multiple Integrals* der findes på Brightspace (Matematiske redskaber).
- Forelæsning 14 (12. oktober): Vi afslutter noterne *Multiple Integrals* og starter med Kap. 5.2, hvor jeg forventer at komme til Afsnit 5.2.1.

Bemærkning vdr. gammafordelingen: Følgende egenskaber for gammafordelingen kan ofte være nyttige:

Sætning B. (Egenskaber af gammafordelingen)

1. Antag at $X \sim Gamma(\alpha_1, \lambda)$ og $Y \sim Gamma(\alpha_2, \lambda)$ er uafhængige. Så gælder der at $X + Y \sim Gamma(\alpha_1 + \alpha_2, \lambda)$.
2. Hvis $X \sim Gamma(\alpha, \lambda)$ og $b > 0$, så er $bX \sim Gamma(\alpha, \lambda/b)$.

Bemærk følgende anvendelse af Sætning B: Hvis X og Y er $Exponential(\lambda)$-fordelte og uafhængige, så er $X + Y \sim Gamma(2, \lambda)$. Dette følger direkte af $Exponential(\lambda) = Gamma(1, \lambda)$ og Sætning B.1 med $\alpha_1 = \alpha_2 = 1$.

Bemærkninger vdr. uafhængighed:

1. To *generelle* stokastiske variable X og Y siges at være uafhængige hvis deres simultane fordelingesfunktion splitter op i produktet af de marginale fordelingsfunktioner, dvs.

$$F_{X,Y}(x, y) = F_X(x)F_Y(y) \quad \text{for alle } x, y \in \mathbb{R}^2.$$

Dette er ækvivalent med

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B) \quad \text{for alle } A, B \subset \mathbb{R}^2,$$

og med

$$E[f(X)g(Y)] = E[f(X)]E[g(Y)]$$

for alle $f : \mathbb{R} \to \mathbb{R}$ og $g : \mathbb{R} \to \mathbb{R}$, så længe middelværdierne er veldefinerede.

2. Lad X og Y betegne *uafhængige* stokastiske variable der har marginale tæthedsfunktioner hhv. f_X og f_Y. Så har X og Y en simultan tæthed givet ved

$$f_{X,Y}(x,y) = f_X(x)f_Y(y) \qquad \text{for alle } x, y \in \mathbb{R}.$$

3. Lad X og Y betegne stokastiske variable med simultan tæthedsfunktion $f_{X,Y}$, der er på formen $f_{X,Y}(x,y) = g(x)h(y)$ for to funktioner $g : \mathbb{R} \to [0,\infty)$ og $h : \mathbb{R} \to [0,\infty)$. Så er X og Y uafhængige og $f_X(x) = cg(x)$ og $f_Y(y) = ch(y)$ for en konstant $c > 0$.

   Så når man skal tjekke om to stokastiske variable er uafhængige, er det nok at tjekke om $f_{X,Y}$ kan skrives som et produkt af to funktioner g og h (man behøver ikke at vide om g og h er tætheder).

Teoretiske øvelser i Uge 8 (23. oktober – 29. oktober): Alle opgaver kan løses efter forelæsningen tirsdag d. 10. oktober.

Del 1:

1. Afsnit 5.4: Øvelserne 1, 2, 4.

2. Eksamen, Vinter 2022/2023, Opgave 2.

Del 2:

1. Afsnit 5.4: Øvelse 3.

2. Afsnit 5.4: Øvelse 11. (b) og (c)

   Hint: I (b) kan man vise og bruge at $P_{X,Y}(x,y) = P_{X|N}(x|x+y)P_N(x+y)$ for $0 \leq x+y \leq 3$.

3. Afsnit 5.4: Øvelse 6.

   Hint: Vi kan repræsentere Y som $Y = \sum_{j=1}^{X} U_j$, hvor $U_1, U_2, \ldots$ er uafhængige stokastiske variable med $U_j \sim Bernoulli(p)$. Det kan bemærkes at stokastiske variable med en repræsentation som Y kaldes **compound Poisson** stokastiske variable og disse spiller en vigtig role i blandt andet matematisk finansiering og forsikringsmatematik.

Afeveringsopgave 8: Afsnit 5.4: Øvelse 9. Tegn også mængden C.

Ugens udfordring: Antag at $X \sim Gamma\big(m, \frac{p}{1-p}\big)$, hvor $m \in \mathbb{N}$ og $p \in (0,1)$ er givne parametre. Givet $X = x$ antages der at Y er en diskret stokastisk variabel med $Y \sim Poisson(x)$.

(a) Vis at
$$Y + m \sim Pascal(m, p).$$
Hint: du må bruge versionen af loven af total sandsynlighed på side 275 i IPSR (formel (5.16)) med $A = \{Y + m = k\}$, $k \in \{m, m+1, \dots\}$.

(b) Hvis $Z \sim Poisson\big(m\frac{1-p}{p}\big)$ vis at $E[Z] = E[Y]$ og $\mathrm{Var}(Z) < \mathrm{Var}(Y)$.

Opgavedel (a) udtrykkes tit ved at sige at Pascalfordelingen er en blanding ('mixture') af Poissonfordelinger. I (b) ser vi at 'mixture' forøger variansen. Derfor bruges Pascalfordlingen (prametriseret lige som i `R`, hvor man tæller antal fiaskoer intil $m$te succes) nogle gange som model på data med udfald i $\{0, 1, 2, \dots\}$, hvis deres varians er større end middelværdien.