Forside

Deleksamen i STK1100 - Sannsynlighetsregning og statistisk modellering.
Onsdag 21. mars 2018
Kl. 09.00 - 11.00 (to timer)

Tillatte hjelpemidler: Godkjent lommeregner og Formelsamling for STK1100 og STK1110. Formelsamlingen er lagt ved oppgave 1.

Oppgavesettet består av 20 flervalgsoppgaver som alle teller like mye. Hvis svarene er gitt som desimaltall, er de rundet av til det gitte antall desimaler.

1 Denne midtveiseksamenen har 20 oppgaver med 5 mulige svar for hver oppgave. På hvor mange måter kan du svare på oppgavesettet?

Velg ett alternativ

- 100
- 3 200 000
- $9.54 \cdot 10^{13}$ ✔
- $2.43 \cdot 10^{18}$
- 15 504

Maks poeng: 1

2 Noora har laget en spilleliste på Spotify med sine ti favorittlåter.
Hun stiller inn slik at låtene blir spilt i tilfeldig rekkefølge.
Hvor mange forskjellige rekkefølger kan de spilles i?

Velg ett alternativ

- 3 628 800 ✔
- 470 001 600
- 5 543 200
- 10 000 000 000
- 362 880

Maks poeng: 1

3 Vi har et stokastisk forsøk med begivenhetene *A* og *B*.
Du får vite at $P(A) = 0.45,\ P(B) = 0.50$ og $P(A \cup B) = 0.80$.
Da er $P(A \cap B)$ lik:

Velg ett alternativ

- 0.15 ✔
- 0.30
- 0.05
- 0.10
- 0.35

Maks poeng: 1

4 Isak har 12 norske og 8 engelske bøker i bokhylla.
Han velger tilfeldig to bøker fra hylla.
Hva er sannsynligheten for at han får én bok i hvert språk?

Velg ett alternativ

- 50.5 % ✔
- 48.0 %
- 50.0 %
- 39.7 %
- 24.0 %

Maks poeng: 1

5 I en eske er det to pengestykker. Det ene pengestykket er normalt, mens det andre har mynt på begge sider. Du velger tilfeldig ett pengestykke og kaster det tre ganger. Hva er sannsynligheten for at du får mynt i alle de tre kastene?

Velg ett alternativ:

- $\frac{11}{16}$
- $\frac{3}{4}$
- $\frac{3}{8}$
- $\frac{27}{64}$
- $\frac{9}{16}$ ✔

Maks poeng: 1

6 I en eske er det to pengestykker. Det ene pengestykket er normalt, mens det andre har mynt på begge sider. Du velger tilfeldig ett pengestykke og kaster det tre ganger. Tenk deg at du får mynt i alle tre kastene. Hva er sannsynligheten for at du har kastet med det normale pengestykket?

Velg ett alternativ

- $\frac{1}{8}$
- $\frac{1}{6}$
- $\frac{1}{9}$ ✔
- $\frac{1}{3}$
- $\frac{3}{16}$

Maks poeng: 1

7 Du kaster en terning gjentatte ganger. Hvor mange ganger må du kaste for at det skal være omtrent $80\,\%$ sannsynlighet for at du får minst én sekser?

Velg ett alternativ

- 7
- 9 ✔
- 8
- 6
- 11

Maks poeng: 1

8 Punktsannsynligheten til den stokastiske variabelen X er gitt ved tabellen:

x	0	1	2	3	4	5
$P(X = x)$	0.10	0.15	0.20	0.25	0.20	0.10

Da er $P(1 \leq X < 4)$ lik:

Velg ett alternativ

- 0.45
- 0.60 ✔
- 0.80
- 0.70
- 0.65

Maks poeng: 1

9 Den kumulative fordelingen til en diskret stokastisk variabel X er gitt ved:

$$F(x) = \begin{cases} 0 & \text{for } x < 1 \\ 0.4 & \text{for } 1 \leq x < 2 \\ 0.7 & \text{for } 2 \leq x < 3 \\ 0.9 & \text{for } 3 \leq x < 4 \\ 1 & \text{for } x \geq 4 \end{cases}$$

Da er forventningen $E(X)$ lik:

Velg ett alternativ

- 1.0
- 2.0 ✔
- 8.5
- 2.5
- 3.0

Maks poeng: 1

10 Denne midtveiseksamenen har 20 oppgaver med 5 mulige svar for hver oppgave. Tenk deg at en student ikke kan svare på noen av spørsmålene og krysser av helt tilfeldig for hver oppgave. Hva er sannsynligheten for at han får minst to riktige svar?

Velg ett alternativ

- 14.0 %
- 53.6 %
- 75.3 %
- 86.3 %
- 93.1 % ✔

Maks poeng: 1

11 SMS-meldinger blir mottatt ved en basestasjon i henhold til en Poissonprosess med rate $\alpha = 15$ per minutt. Hva er sannsynligheten for at det blir mottatt minst to SMS-meldinger i løpet av 10 sekunder?

Velg ett alternativ

- 71.3 % ✔
- 45.6 %
- 54.4 %
- 25.7 %
- 63.5 %

Maks poeng: 1

12 Du kaster en terning gjentatte ganger.
Sannsynligheten for at du får den andre sekseren i det sjuende kastet er lik:

Velg ett alternativ

- $6.7\,\%$ ✔
- $1.1\,\%$
- $23.4\,\%$
- $12.4\,\%$
- $7.8\,\%$

Maks poeng: 1

13 Den stokastiske variabelen X er binomisk fordelt med $n = 100$ og $p = 0.25$.
Videre er $Y = 2X + 3$. Da er standardavviket til Y lik:

Velg ett alternativ

- 17.32
- 11.66
- 18.75
- 4.33
- 8.66 ✔

Maks poeng: 1

14 Den stokastiske variabelen X har momentgenererende funksjon $M_X(t) = \frac{1}{(4-3e^t)^2}$.

Da er forventningen $E(X)$ lik:

Velg ett alternativ

- 6 ✔
- 8
- 2
- 1
- 4

Maks poeng: 1

15 Den stokastiske variabelen X har sannsynlighetstetthet

$$f(x) = \begin{cases} 12x^2(1-x) & \text{for } 0 \leq x \leq 1 \\ 0 & \text{ellers} \end{cases}$$

Da er $P(X \leq 1/2)$ lik:

Velg ett alternativ

- 0.41
- 0.31 ✔
- 0.69
- 0.50
- 0.25

Maks poeng: 1

16 Den stokastiske variabelen X har sannsynlighetstetthet

$$f(x) = \begin{cases} 3x^2 & \text{for } 0 \leq x \leq 1 \\ 0 & \text{ellers} \end{cases}$$

Da er standardavviket til X lik:

Velg ett alternativ

- 0.038
- 0.600
- 0.563
- 0.289
- 0.194 ✔

Maks poeng: 1

17 Den stokastiske variabelen X har sannsynstettstetthet

$$f(x) = \begin{cases} 1 & \text{for } 0 \leq x \leq 1 \\ 0 & \text{ellers} \end{cases}$$

Da er den momentgenererende funksjonen $M_X(t)$ lik:

Velg ett alternativ

- $\frac{e^x - 1}{x}$
- $\frac{e^t - 1}{t}$ ✔
- $\frac{1}{(1-t)^2}$
- $\frac{e^t}{1-t}$
- $\frac{(e^t-1)^2}{t^2}$

Maks poeng: 1

18 (I denne oppgaven får du bruk for tabellen over standardnormalfordelingen som er gitt til venstre for oppgaven.)

Vekten til nyfødte gutter er normalfordelt med forventning $\mu = 3.6$ kg og standardavvik $\sigma = 0.5$ kg. Hva er sannsynligheten for at en nyfødt gutt veier mellom 3.0 kg og 4.0 kg ?

Velg ett alternativ

- 68.3 %
- 50.0 %
- 84.1 %
- 67.3 % ✔
- 78.8 %

Maks poeng: 1

19 Den stokastiske variabelen *X* har kumulativ fordeling

$$F(x) = \begin{cases} 1 - e^{-(x/3)^2} & \text{for } x > 0 \\ 0 & \text{ellers} \end{cases}$$

Da er medianen i fordelingen til X lik:

Velg ett alternativ

- $\frac{1}{2}$
- $3\sqrt{\ln 2}$ ✔
- $\frac{3}{\sqrt{2}}$
- $\frac{3}{2}\ln 2$
- $2\sqrt{\ln 3}$

Maks poeng: 1

20 Den stokastiske variabelen *X* har sannsynlighetstetthet

$$f(x) = \begin{cases} \frac{1}{54}x^2e^{-x/3} & \text{for } x > 0 \\ 0 & \text{ellers} \end{cases}$$

Da er $E(1/X^2)$ lik:

Velg ett alternativ

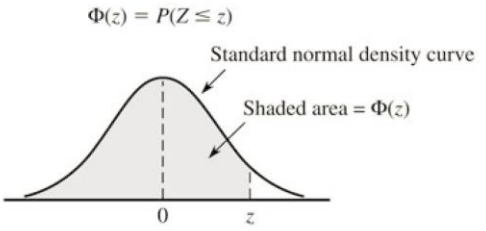
- $\frac{1}{81}$
- $\frac{1}{18}$ ✔
- $\frac{1}{36}$
- $\frac{1}{108}$
- $\frac{1}{54}$

Maks poeng: 1

Question 18
Attached



Table A.3 Standard Normal Curve Areas



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3482
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

(*continued*)

Table A.3 Standard Normal Curve Areas (*cont.*) $\Phi(z) = P(Z \leq z)$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9278	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Question 1
Attached



FORMELSAMLING TIL STK1100 OG STK1110

(Versjon av 14. april 2012)

1. Sannsynlighet

La $A, B, A_1, A_2, \ldots, B_1, B_2, \ldots$ være begivenheter, dvs. delmengder av et utfallsrom Ω.

a) Aksiomene:
Et sannsynlighetsmål P er en funksjon fra delmengder av utfallsrommet Ω til de reelle tall som tilfredsstiller

$$
\begin{aligned}
&P(\Omega) = 1 \\
&P(A) \geq 0 \\
&P(A_1 \cup A_2) = P(A_1) + P(A_2) \quad \text{hvis } A_1 \cap A_2 = \emptyset \\
&P\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} P(A_i) \quad \text{hvis } A_i \cap A_j = \emptyset \text{ for } i \neq j
\end{aligned}
$$

b) $P(A') = 1 - P(A)$

c) $P(\emptyset) = 0$

d) $A \subset B \Rightarrow P(A) \leq P(B)$

e) Addisjonsetningen:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

f) Betinget sannsynlighet:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{hvis } P(B) > 0$$

g) Total sannsynlighet:

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i) \quad \text{hvis } \bigcup_{i=1}^{n} B_i = \Omega \text{ og } B_i \cap B_j = \emptyset \text{ for } i \neq j$$

h) Bayes' setning:

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)} \quad \text{under samme betingelser som i g)}$$

i) A og B er uavhengige begivenheter hvis $P(A \cap B) = P(A)P(B)$

j) $A_1, \ldots, A_n$ er uavhengige begivenheter dersom

$$P(A_{i_1} \cap \cdots \cap A_{i_m}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_m})$$

for alle delmengder av indekser $i_1, i_2, \ldots, i_m$

k) Produktsetningen:

$$\begin{aligned} & P(A_1 \cap \cdots \cap A_n) \\ &= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \end{aligned}$$

2. Kombinatorikk

a) To operasjoner som kan gjøres på henholdsvis n og m måter kan kombineres på $n \cdot m$ måter.

b) Antall ordnete utvalg med tilbakelegging av r elementer fra en mengde med n elementer er n^r

c) Antall ordnete utvalg uten tilbakelegging av r elementer fra en mengde med n elementer er $n(n-1)\cdots(n-r+1)$

d) Antall måter n elementer kan ordnes i rekkefølge på (permuteres) er $n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$

e) Antall ikke-ordete utvalg av r elementer fra en mengde med n elementer er

$$\binom{n}{r} = \frac{n(n-1)\cdots(n-r+1)}{r!} = \frac{n!}{r!\,(n-r)!}$$

f) Antall måter en mengde med n elementer kan deles inn i r delmengder med n_i elementer i den i-te delmengden er

$$\binom{n}{n_1\ n_2\ \cdots\ n_r} = \frac{n!}{n_1!\,n_2!\cdots n_r!}$$

3. Sannsynlighetsfordelinger

a) For en stokastisk variabel X (diskret eller kontinuerlig) er den kumulative fordelingsfunksjonen $F(x) = P(X \leq x)$

b) For en diskret stokastisk variabel X som kan anta verdiene $x_1, x_2, x_3, \ldots$ har vi

$$p(x_j) = P(X = x_j)$$
$$F(x) = \sum_{x_j \leq x} p(x_j)$$

Betingelsene for at $p(x_j)$ skal være en punktsannsynlighet er

$$p(x_j) \geq 0 \quad \text{for alle } j$$
$$\sum_j p(x_j) = 1$$

c) For en kontinuerlig stokastisk variabel X har vi

$$P(a < X < b) = \int_a^b f(x)dx$$
$$F(x) = \int_{-\infty}^{x} f(u)du$$
$$f(x) = F'(x)$$

Betingelsene for at $f(x)$ skal være en sannsynlighetstetthet er

$$f(x) \geq 0$$
$$\int_{-\infty}^{\infty} f(x)dx = 1$$

d) For to stokastiske variabler X og Y (diskrete eller kontinuerlige) er den simultane kumulative fordelingsfunksjonen $F(x, y) = P(X \leq x, Y \leq y)$

e) For diskrete stokastiske variabler X og Y som kan anta henholdsvis verdiene $x_1, x_2, \ldots$ og $y_1, y_2, \ldots$ har vi

$$p(x_i, y_j) = P(X = x_i, Y = y_j)$$
$$F(x, y) = \sum_{x_i \leq x} \sum_{y_j \leq y} p(x_i, y_j)$$

Betingelsene for at $p(x_i, y_j)$ skal være en simultan punktsannsynlighet er analoge til betingelsene i b)

f) For kontinuerlige stokastiske variabler X og Y har vi

$$P\,(\,(X,Y) \in A\,) = \int\int_A f(u,v)dvdu$$

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)dvdu$$

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$$

Betingelsene for at $f(x,y)$ skal være en simultan sannsynlighetstetthet er analoge til betingelsene i c)

g) Marginale punktsannsynligheter:

$$p_X(x_i) = \sum_j p(x_i,y_j) \qquad \text{(for } X\text{)}$$

$$p_Y(y_j) = \sum_i p(x_i,y_j) \qquad \text{(for } Y\text{)}$$

h) Marginale sannsynlighetstettheter:

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy \qquad \text{(for } X\text{)}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx \qquad \text{(for } Y\text{)}$$

i) Uavhengighet:

De stokastiske variablene X og Y er uavhengige dersom

$$p(x_i,y_j) = p_X(x_i)p_Y(y_j) \qquad \text{(diskret)}$$
$$f(x,y) = f_X(x)f_Y(y) \qquad \text{(kontinuerlig)}$$

j) Betingete punktsannsynligheter:

$$p_{X|Y}(x_i|y_j) = \frac{p(x_i,y_j)}{p_Y(y_j)} \qquad \text{(for } X \text{ gitt } Y = y_j\text{)}$$

$$p_{Y|X}(y_j|x_i) = \frac{p(x_i,y_j)}{p_X(x_i)} \qquad \text{(for } Y \text{ gitt } X = x_i\text{)}$$

Det forutsettes at $p_Y(y_j) > 0$ og $p_X(x_i) > 0$, henholdsvis. De betingete punktsannsynlighetene kan behandles som vanlige punktsannsynligheter.

k) Betingete sannsynlighetstettheter:

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)} \qquad \text{(for } X \text{ gitt } Y = y\text{)}$$

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} \qquad \text{(for } Y \text{ gitt } X = x\text{)}$$

Det forutsettes at $f_Y(y) > 0$ og $f_X(x) > 0$, henholdsvis. De betingete sannsynlighetstetthetene kan behandles som vanlige sannsynlighetstettheter.

4. Forventning

a) Forventningsverdien til en stokastisk variabel X er definert ved

$$\mathrm{E}(X) = \sum_j x_j p(x_j) \qquad \text{(diskret)}$$

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x) dx \qquad \text{(kontinuerlig)}$$

b) For en reell funksjon $g(X)$ av en stokastisk variabel X er

$$\mathrm{E}[g(X)] = \sum_j g(x_j) p(x_j) \qquad \text{(diskret)}$$

$$\mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx \qquad \text{(kontinuerlig)}$$

c) $\mathrm{E}(a + bX) = a + b\mathrm{E}(X)$

d) For en reell funksjon $g(X,Y)$ av to stokastiske variabler X og Y er

$$\mathrm{E}\big[g(X,Y)\big] = \sum_i \sum_j g(x_i, y_j) p(x_i, y_j) \qquad \text{(diskret)}$$

$$\mathrm{E}\big[g(X,Y)\big] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f(x,y) dy dx \qquad \text{(kontinuerlig)}$$

e) Hvis X og Y er uavhengige er $\mathrm{E}\big[g(X)h(Y)\big] = \mathrm{E}\big[g(X)\big] \cdot \mathrm{E}\big[h(Y)\big]$

f) Hvis X og Y er uavhengige er $\mathrm{E}(XY) = \mathrm{E}(X) \cdot \mathrm{E}(Y)$

g) $\mathrm{E}\left(a + \sum_{i=1}^{n} b_i X_i\right) = a + \sum_{i=1}^{n} b_i \mathrm{E}(X_i)$

h) Betinget forventning:

$$\mathrm{E}(Y|X=x_i) = \sum_j y_j p_{Y|X}(y_j|x_i) \qquad \text{(diskret)}$$

$$\mathrm{E}(Y|X=x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)dy \qquad \text{(kontinuerlig)}$$

5. Varians og standardavvik

a) Variansen og standardavviket til en stokastisk variabel X er definert ved

$$\mathrm{V}(X) = \mathrm{E}[(X-\mu)^2]$$
$$\mathrm{sd}(X) = \sqrt{\mathrm{V}(X)}$$

b) $\mathrm{V}(X) = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2$

c) $\mathrm{V}(a+bX) = b^2\,\mathrm{V}(X)$

d) Hvis $X_1, \ldots, X_n$ er uavhengige har vi

$$\mathrm{V}\left(a + \sum_{i=1}^n b_i X_i\right) = \sum_{i=1}^n {b_i}^2\,\mathrm{V}(X_i)$$

e)

$$\mathrm{V}\left(a + \sum_{i=1}^n b_i X_i\right) = \sum_{i=1}^n {b_i}^2\,\mathrm{V}(X_i) + \sum_{i=1}^n \sum_{j\neq i} b_i b_j \mathrm{Cov}(X_i, X_j)$$

f) Chebyshev's ulikhet:

La X være en stokastisk variabel med $\mu = \mathrm{E}(X)$ og $\sigma^2 = \mathrm{V}(X)$.
For alle $t > 0$ har vi

$$P(|X-\mu| > t) \leq \frac{\sigma^2}{t^2}$$

6. Kovarians og korrelasjon

a) La X og Y være stokastiske variabler med $\mu_X = \mathrm{E}(X)$, $\sigma_X^2 = \mathrm{V}(X)$, $\mu_Y = \mathrm{E}(Y)$ og $\sigma_Y^2 = \mathrm{V}(Y)$. Da er kovariansen og korrelasjonen til X og Y definert ved

$$\mathrm{Cov}(X,Y) = \mathrm{E}\big[(X-\mu_X)(Y-\mu_Y)\big]$$
$$\rho = \mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

b) $\mathrm{Cov}(X, X) = \mathrm{V}(X)$

c) $\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$

d) X, Y uavhengige $\Rightarrow \mathrm{Cov}(X, Y) = 0$

e)
$$\mathrm{Cov}\left(a + \sum_{i=1}^{n} b_i X_i\,,\; c + \sum_{j=1}^{m} d_j Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} b_i d_j \mathrm{Cov}(X_i, Y_j)$$

f) $-1 \leq \mathrm{Corr}(X, Y) \leq 1$ og $\mathrm{Corr}(X, Y) = \pm 1$ hvis og bare hvis det finnes to tall a, b slik at $Y = a + bX$ (bortsett, eventuelt, på et område med sannsynlighet 0)

7. Momentgenererende funksjoner

a) For en stokastisk variabel X (diskret eller kontinuerlig) er den momentgenererende funksjonen $M_X(t) = \mathrm{E}(e^{tX})$

b) Hvis den momentgenererende funksjonen $M_X(t)$ eksisterer for t i et åpent intervall som inneholder null, så bestemmer den entydig fordelingen til X

c) Hvis den momentgenererende funksjonen $M_X(t)$ eksisterer for t i et åpent intervall som inneholder null, så eksisterer alle momenter til X, og vi kan finne det r-te momentet ved $\mathrm{E}(X^r) = M_X^{(r)}(0)$

d) $M_{a+bX}(t) = e^{at} M_X(bt)$

e) Hvis X og Y er uavhengige er $M_{X+Y}(t) = M_X(t) M_Y(t)$

8. Noen diskrete sannsynlighetsfordelinger

a) Binomisk fordeling:

Punktsannsynlighet: $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad k = 0, 1, \ldots, n$

Momentgenererende funksjon : $M_X(t) = (1 - p + pe^t)^n$

Forventning: $\mathrm{E}(X) = np$

Varians : $\mathrm{V}(X) = np(1-p)$

Tilnærmelse 1: $Z = \dfrac{X - np}{\sqrt{np(1-p)}}$ er tilnærmet standard normalfordelt når np og $n(1-p)$ begge er tilstrekkelig store (minst 10)

Tilnærmelse 2: X er tilnærmet Poisson fordelt med parameter $\lambda = np$ når n er stor og p er liten

Addisjonsregel: $X \sim$ binomisk (n,p), $Y \sim$binomisk (m,p)
og X, Y uavhengige $\Rightarrow X+Y \sim$ binomisk $(n+m,p)$

b) Geometrisk fordeling:

Punktsannsynlighet: $P(X=k) = (1-p)^{k-1}p \qquad k = 1,2,\ldots$

Momentgenererende funksjon : $M_X(t) = e^t p/[1-(1-p)e^t]$

Forventning: $\mathrm{E}(X) = 1/p$

Varians: $\mathrm{V}(x) = (1-p)/p^2$

Addisjonsregel: Hvis X er geometrisk fordelt med sannsynlighet p så er $X-1$ negativt binomisk $(1,p)$. Derfor hvis X og Y er geometrisk fordelte med samme p og uavhengige så er $X+Y-2$ negativt binomisk $(2,p)$

c) Negativ binomisk fordeling:

Punktsannsynlighet: $P(X=k) = \binom{k+r-1}{r-1}p^r(1-p)^k \qquad k = 0,1,2,\ldots$

Momentgenererende funksjon : $M_X(t) = \{p/[1-(1-p)e^t]\}^r$

Forventning: $\mathrm{E}(X) = r(1-p)/p$

Varians: $\mathrm{V}(X) = r(1-p)/p^2$

Addisjonsregel: $X \sim$ negativ binomisk (r_1,p),
$Y \sim$ negativt binomisk (r_2,p)
og X, Y uavhengige
$\Rightarrow X+Y \sim$ negativt binomisk (r_1+r_2,p)

d) Hypergeometrisk fordeling:

Punktsannsynlighet: $P(X=k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$

Forventning: $\mathrm{E}(X) = n \cdot \frac{M}{N}$

Varians: $\mathrm{V}(X) = n\frac{M}{N}(1-\frac{M}{N})\frac{N-n}{N-1}$

Tilnærmelse: X er tilnærmet binomisk $(n, \frac{M}{N})$ når n er mye mindre enn N

e) Poisson fordelingen:

Punktsannsynlighet: $P(X=k) = \frac{\lambda^k}{k!}e^{-\lambda} \qquad k = 0,1,\ldots$

Momentgenererende funksjon : $M_X(t) = e^{\lambda(e^t-1)}$

Forventning: $\mathrm{E}(X) = \lambda$

Varians: $\mathrm{V}(X) = \lambda$

Tilnærmelse: $Z = \dfrac{X - \lambda}{\sqrt{\lambda}}$ er tilnærmet standard normalfordelt når λ er tilstrekkelig stor (minst 10)

Addisjonsregel: $X \sim$Poisson (λ_1), $Y \sim$Poisson (λ_2) og X, Y uavhengige $\Rightarrow X + Y \sim$ Poisson $(\lambda_1 + \lambda_2)$

e) Multinomisk fordeling:

Punktsannsynlighet: $P(N_1 = n_1, \ldots, N_r = n_r) = \frac{n!}{n_1! \cdots n_r!} p_1^{n_1} \cdots p_r^{n_r}$
Her er $\sum_{i=1}^{r} p_i = 1$ og $\sum_{i=1}^{r} n_i = n$

Marginalfordeling: $N_i \sim$ binomisk (n, p_i)

9. Noen kontinuerlige sannsynlighetsfordelinger

a) Normalfordelingen:

Tetthet: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$ $\quad -\infty < x < \infty$

Momentgenererende funksjon : $M_X(t) = e^{\mu t} e^{\sigma^2 t^2/2}$

Forventning: $\mathrm{E}(X) = \mu$

Varians: $\mathrm{V}(X) = \sigma^2$

Transformasjon: $X \sim N(\mu, \sigma^2) \Rightarrow a + bX \sim N(a + b\mu, b^2\sigma^2)$
$X \sim N(\mu, \sigma^2) \Rightarrow Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$

Addisjonsregel: $X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$, X, Y uavhengige
$\Rightarrow X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$

b) Eksponentialfordelingen:

Tetthet: $f(x) = \lambda e^{-\lambda x}$ $\quad x > 0$

Momentgenererende funksjon : $M_X(t) = \lambda/(\lambda - t)$ for $t < \lambda$

Forventning: $\mathrm{E}(X) = 1/\lambda$

Varians: $\mathrm{V}(X) = 1/\lambda^2$

Addisjonsregel: $X \sim \exp(\lambda)$, $Y \sim \exp(\lambda)$, X og Y uavhengige
$\Rightarrow X + Y \sim \mathrm{gamma}(2, 1/\lambda)$

c) Gammafordelingen:

Tetthet: $f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}$ $\quad x > 0$

Gammafunksjonen: $\Gamma(\alpha) = \int_0^\infty u^{\alpha-1}e^{-u}du$
$\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$
$\Gamma(n) = (n-1)!$ når n er et helt tall
$\Gamma(1/2) = \sqrt{\pi}, \quad \Gamma(1) = 1$

Momentgenererende funksjon : $M_X(t) = [1/(1-\beta t)]^\alpha$

Forventning: $\mathrm{E}(X) = \alpha\beta$

Varians: $\mathrm{V}(X) = \alpha\beta^2$

Addisjonsregel: $X \sim \text{gamma}(\alpha, \beta), \quad Y \sim \text{gamma}(\delta, \beta)$,
X og Y uavhengige $\Rightarrow X + Y \sim \text{gamma}(\alpha + \delta, \beta)$

d) Kji-kvadratfordelingen:

Tetthet: $f(v) = \frac{1}{2^{n/2}\Gamma(n/2)} v^{(n/2)-1}e^{-v/2} \qquad v > 0$
n er antall frihetsgrader

Forventning: $\mathrm{E}(V) = n$

Varians: $\mathrm{V}(V) = 2n$

Addisjonsregel: $U \sim \chi^2_n$, $V \sim \chi^2_m$, U og V uavhengige
$\Rightarrow U + V \sim \chi^2_{n+m}$

Resultat: $Z \sim N(0,1) \Rightarrow Z^2 \sim \chi^2_1$

e) Students t-fordeling:

Tetthet: $f(t) = \frac{\Gamma[(n+1)/2]}{\sqrt{n\pi}\Gamma(n/2)}(1 + \frac{t^2}{n})^{-(n+1)/2} \qquad -\infty < t < \infty$
n er antall frihetsgrader

Forventning: $\mathrm{E}(T) = 0 \qquad (n \geq 2)$

Varians: $\mathrm{V}(T) = n/(n-2) \qquad (n \geq 3)$

Resultat: $Z \sim N(0,1), \quad U \sim \chi^2_n, \quad Z, U$ uavhengige $\Rightarrow Z/\sqrt{U/n} \sim t_n$

f) Binormal fordeling:

Tetthet:

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right\}$$

Marginalfordeling: $X \sim N(\mu_X, \sigma_X^2), \quad Y \sim N(\mu_Y, \sigma_Y^2)$

Korrelasjon: $\mathrm{Corr}(X,Y) = \rho$

Betinget fordeling: Gitt $X = x$ er Y normalfordelt med
forventning $\mathrm{E}(Y|X=x) = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X)$
og varians $\mathrm{V}(Y|X=x) = \sigma_Y^2(1-\rho^2)$

10. Ett normalfordelt utvalg

Hvis $X_1, X_2, \ldots, X_n$ er uavhengige og $N(\mu, \sigma^2)$-fordelte så har vi at:

a) $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{og} \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ er uavhengige

b) $\overline{X} \sim N(\mu, \sigma^2/n)$

c) $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$

d) $\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$

11. To normalfordelte utvalg

La $X_1, X_2, \ldots, X_n$ være uavhengige og $N(\mu_X, \sigma^2)$-fordelte, og $Y_1, Y_2, \ldots, Y_m$ uavhengige og $N(\mu_Y, \sigma^2)$-fordelte. De to utvalgene er uavhengige av hverandre. La $\overline{X}$, $\overline{Y}$, S_X^2 og S_Y^2 være definert i henhold til 10a). Da har vi at:

a) $S_p^2 = [(n-1)S_X^2 + (m-1)S_Y^2]/(m+n-2)$ er en vektet estimator for σ^2

b) $\overline{X} - \overline{Y} \sim N\left(\mu_X - \mu_Y, \sigma^2(\frac{1}{n} + \frac{1}{m})\right)$

c) $(n+m-2)S_p^2/\sigma^2 \sim \chi^2_{m+n-2}$

d) $\frac{\overline{X}-\overline{Y}-(\mu_X-\mu_Y)}{S_p\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim t_{m+n-2}$

12. Regresjonsanalyse

Anta at $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i\,;\ i = 1, 2, \ldots, n\,;$ hvor x_i-ene er kjente tall og ϵ_i-ene er uavhengige og $N(0, \sigma^2)$-fordelte. Da har vi at:

a) Minste kvadraters estimatorer for β_0 og β_1 er

$$\hat{\beta}_0 = \overline{Y} - \hat{\beta}_1\overline{x} \quad \text{og} \quad \hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

b) Estimatorene i a) er normalfordelte og forventningsrette, og

$$\mathrm{Var}(\hat{\beta}_0) = \frac{\sigma^2\sum_{i=1}^{n} x_i^2}{n\sum_{i=1}^{n}(x_i - \overline{x})^2} \quad \text{og} \quad \mathrm{Var}(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

c) La $\mathrm{SSE} = \sum_{i=1}^{n}(Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$. Da er $S^2 = \mathrm{SSE}/(n-2)$ en forventningsrett estimator for σ^2, og $(n-2)S^2/\sigma^2 \sim \chi^2_{n-2}$