Sannolikhetskalkyl 2

Föreläsningsanteckningar

John Möller

Contents

Dag

1

Introduktion

1 John Möllers notationer

Jag kommer alltid ha $\subset$ betyder delmängd (som kan vara likamed). Det har d.v.s en motsvarighet till $\leq$ (och inte $<$).

2 Formalia

Daniel Ahlberg

Kurshemsidan Litteratur: An intermediate course in probability, 2nd Edition, Springer (2009).

- 6 Stycken frivilliga inlämmningsuppgifter, som ger bonuspoäng till tentan.
- Muntlig redovisning (mer information skall komma).

Det enda obligatoriska är tentan (men de vill väldigt gärna att man går på muntliga redovisning).

3 Innehåll av kursen

Målet är att få fördjupad förståelse av grundläggande begrepp och metoder.

Speciellt kolla på:

- Flerdimensionella stokastiska variabler, speciellt den multivariata normalfördelningen.
- Tekniker såsom betingning och transformationer.
- Konvergensbregrepp inom sannolikhetsteorin.

4 Föreläsning

Syftet med sannolikhetskalkyl är att skapa en matematisk grund för att analysera modeller för slumpmässiga fenomen

4.1 Några exempel

Exempel (Smittspridning). F/ Vad krävs för att en smtita skall ge upphov till ett stort utbrott?

$$R_0 := \text{ förväntat antal nya fall en person smittad individ ger upphov till }.$$

Om $R_0 > 1$ medför att ett stort utbrott är möjligt. Medan om $R_0 < 1$ medför det att ett stort utbrott är ej möjligt. (förgreningsprocesser)

Exempel (Kortblandningar). Hur många gånger och hur ska man blanda för att man ska vara "nöjd"? Att dela i två är kräver typ bara 7 gånger. Kanske tusen gånger på ett annat sätt.
(markovkedjor)

4.2 Repetition

Det mest grundläggande är ett sannolikhetsrum.

Definition 4.1 (Sannolikhetsrum). Ett sannolikhetsrum är en trippel $(\Omega, \mathcal{F}, \mathcal{P})$.

- Mängden Ω kallas utfallsrummet,
- mängden $\mathcal{F}$ är en samling delmängder av Ω kallade händelser.
- Funktionen $\mathcal{P} : \mathcal{F} \to [0, 1]$

  Funktionen uppfyller följande krav:
- $\mathcal{P}(A) \geq 0$ för alla $A \in \mathcal{F}$
- $\mathcal{P}(\Omega) = 1$
- Om $A_1, A_2, \ldots$ är parvis disjunkta händelser så gäller $\mathcal{P}\left(\bigcup_{k=1}^{\infty}\right) = \sum_{k=1}^{\infty} \mathcal{P}(A_k)$

  Funktionen $\mathcal{P}$ kallas för sannolikhetsmått.

Kommentar (Formellt extrakrav). För formell korrekthet krävs att $\mathcal{F}$ är en σ-algebra. Man har det för att man ska kunna ta snitt och så vidare... (överkurs)

Kommentar (Tolkning). Om man uppreppar så kommer den relativa frekvensen närma sig sannolikheten för händelsen.

4.3 Egenskaper hos $\mathcal{P}$

Från definition härleds att

- $\mathcal{P} \in [0, 1]$ för alla $A \in \mathcal{F}$
- $\mathcal{P}(\emptyset) = 0$
- Om A och B är disjunkta, s är $\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B)$.
- Om $A \subset B$, så $\mathcal{P}(A) \leq \mathcal{P}(B)$

Låt $(\Omega, \mathcal{F}, P)$ vara ett sannolikhetsrum.

4.4 Några fler definitioner

Definition 4.2 (betingning). Låt $A, B \in \mathcal{F}$ vara två händelser och antag att $P(B) > 0$. Den betingade sannolikheten för A givet B defineras som

$$P(A|B) := \frac{P(A \cap B)}{P(B)}.$$

Definition 4.3 (Oberoende). A och B är oberoende om

$$P(A \cap B) = P(A) \cdot P(B).$$

Definition 4.4 (Stokastisk variabel)**.** En stokastisk variabel är en (mätbar) funktion $X : \Omega \to \mathbb{R}$.
Fördelningsfunktion $F_X : \mathbb{R} \to [0, 1]$ av X ges av

$$F_X := P(X \leq x).$$

Och i teorin (John: är inte helt säker på vad det innebär) bestämmer den stokastiska variabeln.

Två klasser av fördelningar. Låt massfunktionen av X vara

$$p_X(x) := P(X = x).$$

Kommentar. $p_X(x) > 0$ för som mest uppräkneligt många värden på X.
(Illustration)

- X har en diskret fördelning om $F_X(x) = \sum_{y \leq x} p_X(x)$ för alla $x \in \mathbb{R}$.
- X har en kontinuerlig fördelning om det existerar en (integrerbar) funktion $f_X : \mathbb{R} \to [0, \infty)$ sådan att $F_x(x) = \int_{-\infty}^{x} f_X(y)dy$ för alla $x \in \mathbb{R}$.

f_X kallas täthetsfunktion behöver inte vara unik (för integraler).

Exempel (Täthetsfunktionen är ej unik)**.** Båda följande exempel ger en täthet för en likformig fördelning.
funktionen

$$f_1(x) = \begin{cases} 1 & \text{om } x \in [0, 1] \\ 0 & \text{annars} \end{cases}.$$

Sen kan vi definera

$$f_2(x) = \begin{cases} 1 & \text{om } x \in [0, 1] \text{ och } x \neq 0.5 \\ 0 & \text{annars} \end{cases}.$$

Definition 4.5 (Väntevärde)**.** Väntevärdet av en stokastisk variabel X defineras som

$$E[X] := \begin{cases} \sum_{x \in \mathbb{R}} x \cdot p_X(x) & \text{om } X \text{ diskret} \\ \int_{-\infty}^{\infty} x f_X(x)dx & \text{om } X \text{ kontinuerlig} \end{cases},$$

givet att dessa är väldefinerade/ändliga.

Väntevärde ger det förväntade värdet man får från en stokastisk variabel.

Kommentar. Eftersom $p_X(x)$ är nollskiljd för ett uppräckneligt antal punkter så kommer summan vara uppräknelig.

Definition 4.6 (Varians)**.** Variansen av X, då $E[X]$ är väldefinerat, ges av följande uttryck

$$\text{Var}(X) := E\left[(X - E[X])^2\right] = E[X^2] - E[X]^2.$$

Uppgift 4.1. Visa att om $\operatorname{Var}(X) = 0$ bör vara ekvivalent med att det existerar ett $c \in \mathbb{R}$ så att:

$$P(X = c) = 1.$$

Låt X och Y vara stokastiska variabler definerade på samma sannolikhetsrum.

Definition 4.7 (Bivariata fördelning)**.** Bivariata fördelningen för X och Y är

$$F_{X,Y}(x, y) := \mathcal{P}(X \leq x, Y \leq y).$$

Definition 4.8 (Oberoende)**.** X och Y är oberoende om

$$F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y),$$

för alla $x, y \in \mathbb{R}$.

Definition 4.9 (Kovarians)**.** Kovariansen av X och Y är

$$\operatorname{Cov}(X, Y) := E\left[(X - E[X])(Y - E[Y])\right] = E[XY] - E[X]E[Y].$$

Om X och Y är oberoende så gäller att

$$E[XY] = E[X]E[Y],$$

så

$$\operatorname{Cov}(X, Y) = 0.$$

Definition 4.10 (Korrelationskoefficienten)**.** Korrelationskoefficienten av X och Y ges av

$$\rho_{X,Y} := \frac{\operatorname{Cov}(X, Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}}.$$

Den är alltid mellan -1 och 1.

Sats 4.1 (Markovs olikhet)**.** Om X är icke-negativt så gäller

$$P(X \geq a) \leq \frac{E[X]}{a}.$$

Sats 4.2 (Chebyshevs olikhet)**.** Om $E[X]$ existerar, så gäller

$$P(|X - E[X]| > a) \leq \frac{\operatorname{Var}(X)}{a^2}.$$

4.5 Tolkning av sannolikhet

Kom ihåg tolkningen av sannolikheter som den relativa frekvensen med vilken en händelse inträffar då ett experiment uppreppas oberoende av varandra många gånger.

> **Fråga.** Varför stämmer detta?

Låt oss betrakta något "experiment" $(\Omega, \mathcal{F}, P)$ och någon händelse A. Låt

$$X_k = \begin{cases} 1 & \text{om } A \text{ inträffar omgång } k \\ 0 & \text{annars} \end{cases}.$$

Dessa variabler är oberoende om experimentet utförs oberoende av tidigare omgångar.

Den relativa frekvensen för A i n omgångar är

$$\frac{1}{n}\sum_{k=1}^{n} X_k.$$

> **Fråga.** Gäller det att $\frac{1}{n}\sum_{k=1}^{n} X_k \approx P(A)$?

Observera att

$$\begin{aligned} E\left[\frac{1}{n}\sum_{k=1}^{n} X_k\right] &= \frac{1}{n}\sum_{k=1}^{n} E[X_n] = E[X_1] \\ &= 0 \cdot P(A^c) + 1 \cdot P(A) = P(A). \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(\frac{1}{n}\sum_{k=1}^{n} X_k) &= \frac{1}{n^2}\sum_{k=1}^{n} Var(X_k) \\ &= \frac{1}{n} P(A)[1 - P(A)]. \end{aligned}$$

Chebyshevs olikhet ger

$$\begin{aligned} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - P(A)\right| > \epsilon\right) &\leq \frac{\mathrm{Var}(\frac{1}{n}\sum_{k=1}^{n} X_k)}{\epsilon^2} \\ &= \frac{P(A)\,[1 - P(A)]}{n \cdot \epsilon^2}. \end{aligned}$$

Detta är litet om n är stort, och blir mindre då n ökar. Detta motiverar tolkningen.

Dag

2

Flerdimensionella stokastiska variabler

Exempel.

Dag

4

Mer om betingade väntevärden

5 Betingning

Om (X,Y) är en diskret slumpvektor så är den betingade fördelningen av Y givet $X=x$ fördelningen för Y med avseende på $P(\star|X=x)$.

Om (X,Y) kontinuerlig så är den betingade fördelningen för Y givet $X=x$ den fördelaning med täthet

$$f_{Y|X=x}(y) := \frac{f_{X,Y}(x,y)}{f_X(x)},$$

för alla $y \in \mathbb{R}$.

6 Motivering, kontinuerliga fallet

Låt X vara en kontinuerlig variabel. Låt också $\Delta > 0$, och sätt $x_k = \Delta \cdot k$ och $I_k = (x_{k-1}, x_k]$.

Eftersom Riemann-integralen är gränsvärdet av en summa, så får vi

$$P\left(X \in (a,b]\right) = \int_a^b f_X(x)dx = \lim_{\Delta \to 0} \sum_k f_X(x_k)\Delta.$$

där summan gör över alla k så att $x_k \in (a,b]$.

Låt A vara en händlese. Ett rimligt försök är att definiera

$$P(A|X=x) = \lim_{\Delta \to 0} P(A|x \in (x-\Delta, x]).$$

Exempel. Antag $A = \{X \geq 0\}$. Ovan definitionen ger då

$$\begin{aligned} P(A|X=0) &= \lim_{\Delta \to 0} P(A|X \in (-\Delta, 0]) \\ &= \lim_{\Delta \to 0} \frac{P(X=0)}{P(X \in (\Delta, 0])} = 0. \end{aligned}$$

Observera dock

$$\lim_{\Delta \to 0} P(A|X \in (0,\Delta]) = \lim_{\Delta \to 0} \frac{P(X \in (0,\Delta])}{P(X \in (0,\Delta])} = 1.$$

Att definerna $P(A|X = x)$ med ett gränsvärde ger alltså problem.

Låt oss anta att

$$P(A|X = x) = \lim_{\Delta \to 0} P(A|X \in (x - \Delta, x])$$

för de flesta $x \in \mathbb{R}$. Notera att

$$\begin{aligned} P(A, X \in (a, b]) &\approx \sum_k P(A, X \in I_k) \\ &= \sum_k P(A|X \in I_k) \cdot P(X \in I_k). \end{aligned}$$

Ungefär så är $P(X \in I_k) \approx f_X(x_k) \cdot \Delta$. Om vi nu skickar $\Delta \to 0$ så får vi att $P(A|X = x)$ borde uppfylla följande ekvation:

$$P(A, X \in (a, b]) = \int_a^b P(A|X = x) f_X(x) dx.$$

(Ovan är ekvation **) Lösning den betingade sannolikheten av A givet X som någon funktion $x \mapsto P(A|X = x)$ som uppfyller ekvationen ovan.

Kommentar. Funktionen $x \mapsto P(A|X = x)$ behöver inte vara unika, men två funktioner som uppfyller ekvationen är lika för nästan alla $x \in \mathbb{R}$.

Då är (X, Y) är en kontinuerlig slumpvektor så är (**) ekvivalent med (sätt $A = \{Y \le y\}$, $a = -\infty$)

$$F_{X,Y}(x, y) = \int_{-\infty}^x F_{Y|X=u} f_x(u) dx.$$

Kommentar. Vår tidigare definition av $F_{Y|X=u}(y)$ ges av

$$F_{Y|X=u}(y) = \int_{-\infty}^y f_{Y|X=u} du.$$

Insatt i HL får vi

$$\begin{aligned} HL &= \int_{-\infty}^x \int_{-\infty}^y \frac{f_{X,Y}(u, v)}{f_X(u)} dv f_X(u) dx \\ &= \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(u, v) du dv \\ &= F_{X,Y}(x, y). \end{aligned}$$

Alltså är vårde definitioner konsistenta med diskussionen ovan.

Exempel. Låt oss anta att antalet röda blodkroppar i ett blodprov följer en Poissonfördelning, men att parametern varierar mellan individer.
Detta kan beskrivas via följande mmodell: Antalet blodkroppar X följer följande samband:

$$X|M = m \sim \text{Po}(m),$$

med $M \sim F$, där F är någon känd eller okänd fördelning. Fördelningen för (X, M)

beskrivs här indirekt. Om F är diskret så får vi

$$P(X=k, M=l) = P(X=k|M=l)\cdot P(M=l).$$

Där $P(X=k|M=l) \sim \mathrm{Po}(m)$ och $P(M=l) \sim F$.
I fallet då F är kontinuerlig så får vi från (**) att

$$P(X=k, M\in(a,b]) = \int_a^b P(X=k|M=m)\cdot f_M(m)dm.$$

Exempel. Antag $M \sim \exp(1)$.

$$\begin{aligned}
P(X=k) &= \int_0^\infty P(X=k|M=m)\cdot e^{-m}dm \\
&= \int_{-0}^\infty e^{-m}\frac{m^k}{k!}\cdot e^{-m}dm \\
&= \frac{1}{2^{k+1}}\int_0^\infty 2^{k+1}\cdot e^{-2m}\cdot\frac{m^k}{k!}dm \\
&= \frac{1}{2^{k+1}}
\end{aligned}$$

Sista steget fås genom att se att integranden är tätheten av en Τ -fördelning
Alltså har X en geometrisk fördelning.

7 Betingad varians

Definition 7.1 (Betingad varians)**.** Låt (X,Y) vara en (kontinuerlig eller diskret) slumpvektor. Den betingade variansen för Y givet $X=x$ är

$$\mathrm{Var}(Y|X=x) := E\left[(y-E[Y|X=x])^2|X=x\right],$$

förutsatt att $E[Y|X=x]$ är väldefinerat.

För varje $x\in\mathbb{R}$ får vi ett värde på funktionen $\mathrm{V}(X) := \mathrm{Var}(Y|X=x)$. Dock är $V(X)$ en stokastisk variabel och vi skriver $V(X) = \mathrm{Var}(Y|X)$.

Sats 7.1. Antag att $E[Y^2]<\infty$. Då gäller

$$\mathrm{Var}(Y) = E\left[\mathrm{Var}(Y|X)\right] + \mathrm{Var}(E\left[Y|X\right]^2).$$

Bevis. Notera att

$$\begin{aligned}
\mathrm{Var}(Y) &= E\left[(Y-E[Y])^2\right] \\
&= E\left[(Y-E[Y|X]+E[Y|X]-E[Y])^2\right] \\
&= E\left[(Y-E[Y|X])^2\right] \\
&+ E\left[E[Y|X]-E[Y]^2\right] \\
&+ 2E\left[(Y-E[Y|X])(E[Y|X]-E[Y])\right].
\end{aligned}$$

Låt de tre sista termerna i kalkulationen vara A, B respektive C.

$$\begin{aligned} A &= E\left[(Y - E[Y|X])^2\right] \\ &= E\left[E\left[(Y - E[Y|X])^2 |X\right]\right] \\ &= E\left[\mathrm{Var}(Y|X)\right]. \end{aligned}$$

$$\begin{aligned} B &= E\left[(E[Y|X] - E[Y])^2\right] \\ &= \mathrm{Var}(E[Y|X]). \end{aligned}$$

$$\begin{aligned} C &= E\left[E\left[(Y - E[Y|X])(E[Y|X] - E[Y])|X\right]\right] \\ &= E\left[(E[Y|X] - E[Y])\left(E[Y|X] - E[Y|X]\right)\right] \\ &= 0. \end{aligned}$$

□

Exempel. En pinne har längd 1 bryts på en lokformigt vald punkt. Upprepa. Låt X beteckna första brytpunkten och Y den andra.
Alltså $Y|X = x \sim \text{likf}[0, x]$ och $X \sim \text{likf}[0, 1]$.
Vi har

$$\begin{aligned} E[Y|X = x] &= \frac{x}{2} \\ \mathrm{Var}(Y|X = x) &= \frac{x^2}{12}. \end{aligned}$$

Vi får att

$$\begin{aligned} \mathrm{Var}(E[Y|X]) &= \mathrm{Var}(\frac{X}{2}) \\ &= \frac{1}{4}\mathrm{Var}(X) \\ &= \frac{1}{4} \cdot \frac{1}{12} \cdot E[\mathrm{Var}(Y|X)] \\ &= \frac{1}{12} \cdot \frac{1}{3}. \end{aligned}$$

Sats ger

$$\mathrm{Var}(Y) = \frac{1}{4} \cdot \frac{1}{12} + \frac{1}{3} \cdot \frac{1}{12} = \frac{7}{144}.$$

Uppgift 7.1. Låt (X, Y) vara en slumpvektor med täthet

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{2} & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ \frac{1}{8} & 1 < x \leq 3, 1 < y \leq 3 \end{cases}.$$

a) Bestäm den betingade tätheten för Y givet $X = x$ för $x \in [0, 3]$ b) Beräkna $E[Y]$
c) Avgör om X och Y är oberoende.

Lösning 7.1.1. a) Vi behöver f_X

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy.$$

Rita upp på en xy graf.

Sats 7.2. Låt $X : \Omega \to \mathbb{R}^n$ vara en kontinuerlig slumpvektor med densitet f_X.
Om $g, g^{-1} \in C^1$. Låt $g : S \to T \subset \mathbb{R}^n$ vara en bijektion och sätt $Y = g(X)$.
Då är $Y : \Omega \to \mathbb{R}^n$ kontinuerlig med densitet

$$f_Y(y) = \begin{cases} f_X(g_1^{-1}(y), \dots, g_n^{-1}(y)) \cdot |J| & y \in T \\ 0 & y \notin T \end{cases},$$

där J är jakobianen för g^{-1}.

Exempel. Låt X och Y vara oberoende $\mathrm{N}(0,1)$-fördelade. Visa att $X+Y$ och $X-Y$ är oberoende $\mathrm{N}(0,2)$-fördelade.

Dag

5

Transformer

Låt $X_1, X_2, \dots, X_n$ vara oberoende och likafördelade.

Låt

$$S_n = X_1 + \cdots + X_n.$$

I kontinuerliga fallet gäller

$$P(X_1 + X_2 \leq x) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f_{X_1}(v) f_{X_2}(u-v) dv du.$$

Genom transformer kan vi ersätta multipelintegralen med multiplikation.

Definition 7.2 (sannolikhetsgenererande funktionen)**.** Låt X vara icke-negativ och heltalsvärd stokastisk variabel.
Den sannolikhetsgenererande funktionen (SGF) av X är

$$g_X(t) := \sum_{n \geq 0} t^n \cdot P(X = n).$$

Kommentar.

$$g_X(0) = P(X = 0),$$
$$g_X(1) = \sum_{n \geq 0} P(X = n) = 1.$$

Kommentar. För $|t| \leq 1$ så

$$\begin{aligned}
&\left| g_X(t) - \sum_{n=0}^{N-1} t^n P(X=n) \right| \\
&\leq \sum_{n=N}^{\infty} |t|^n \cdot P(X=n) \\
&\leq \sum_{n \geq N} P(X=n) \to 0,
\end{aligned}$$

då $N \to \infty$.
Så för åtminstone $|t| \leq 1$ så är g_X väldefinerad.

8 potensserier och satser om SGF

Låt $(a_n)_{n\geq 0}$ vara en följd av rella tal och sätt

$$f(t) := \sum_{n=0}^{\infty} a_n t^n.$$

Lemma 8.1. Om f är väldefinerad (serien konvergerar) för $|t| < 1$ så är f kontinuerlig och deriverbar med

$$f'(t) = \sum_{n\geq 1} n a_n t^{n-1},$$

där f' är väldefinerad för $|t| < 1$.

Lemma 8.2. Om $\sum_{n\geq 0} a_n$ konvergerar, så gäller det att

$$\lim_{t\to 1^-} f(t) = \sum_{n\geq 0} a_n.$$

Bevis. Bevisas i kursen analysens grunder. □

Sats 8.3. Låt X, Y vara icke-negativa och heltalsvärda stokastiska variabler. Då gäller

$$g_X = g_Y \iff p_X = p_Y.$$

Bevis. Beviset av $\leftarrow$ följer av definitionen.
Vi vill visa att g_X bestämmer p_X. Notera först att $g_X(0) = P(X=0)$. Eftersom g_X är väldefinerad för $|t| < 1$ så ger Lemma 1 att

$$g_X'(t) = \sum_{n\geq 1} n \cdot t^{n-1} P(X=n).$$

Eftersom g_X' väldefinerad för $|t| < 1$ kan vi fortsätta derivera. Samtliga ordningens derivata existerar och ges av

$$\begin{aligned}
&g_X^k(t) \\
&= \sum_{n\geq k} n(n-1)(n-2)\dots(n-k+1)t^{n-k} \cdot P(X=n).
\end{aligned}$$

Låt för $t = 0$ får vi

$$g_X^{(k)}(0) = k!P(X = k).$$

Alltså bestämmer g_X $P(X = k)$ för varje heltal $k \geq 0$. □

Sats 8.4. Låt $X_1, X_2, \ldots, X_n$ vara oberoende icke-negativa och heltalsvärda stokastiska variabler. Sätt $S_n = X_1 + \cdots + X_n$. Då gäller

$$g_{S_n}(t) = \prod_{k=1}^{n} g_{X_k}(t).$$

Bevis. Notera att för $t \neq 0$ så gäller

$$g_X(t) \sum_{n \geq 0} t^n P(X = n) = E[t^x].$$

Detta ger

$$\begin{aligned} g_{S_n}(t) &= E[t^{X_1+X_2+\cdots+X_n}] \\ &= E[t^{X_1}]E[t^{X_2}]\ldots E[t^{X_n}] \\ &= \prod_{k=1}^{n} g_{X_k}(t). \end{aligned}$$

(Gäller också för $t = 0$ eftersom kontinuerliga funktioner.) □

Exempel. Låt $X_1, \ldots, X_n$ vara oberoende bern(p)-fördelade. Sätt $S_n = X_1 + \cdots + X_n$. Per definition

$$\begin{aligned} g_{X_1}(t) &= P(X = 0) + tP(X = 1) \\ &= (1 - p) + tp. \end{aligned}$$

Enligt sats

$$\begin{aligned} g_{S_n}(t) &= (g_{X_1}(t))^n \\ &= ((1 - p) + tp)^n \\ &= \sum_{k=1}^{n} \binom{n}{k} (1 - p)^{n-k}(tp)^k. \end{aligned}$$

Låt $Y \sim \text{Bin}(n, p)$. Per definition

$$\begin{aligned} g_Y(t) &= \sum_{k=0}^{n} t^k P(Y = k) \\ &= \sum_{k=0}^{n} t^k \binom{n}{k} p^k (1 - p)^{n-k}. \end{aligned}$$

Alltså är $g_{S_n} = g_Y$, och enligt sats har S_n och Y har samma fördelning.

9 Momentgenererande funktioner

Definition 9.1 (Momentgenererande funktionen)**.** Den momentgenererande funktionen av en stokastisk variabel X ges av

$$\Psi_X(t) := E[e^{tX}],$$

givet att det existerar något $h > 0$ så att väntevärdet ovan är ändligt för $|t| \leq h$.

Sats 9.1. Låt X vara en stokastisk variabel för vilken $\Psi_X(t)$ existerar för $|t| < h$ för något $h > 0$. Då gäller

(a) $E[|X|^r] < \infty$ för alla $r > 0$

(b) $\Psi_X^{(n)}(0) = E[X^n]$ för $n \geq 1$.

Bevis. (a): Tag $0 < t < h$. Då gäller

$$\begin{aligned} E[e^{|tX|}] &\leq E[e^{tX} + e^{-tX}] \\ &= \Psi_X(t) + \Psi_X(-t) < \infty. \end{aligned}$$

Funktionen e^{tx} växer snabbare än x^r. Det vill säga för varje $t > 0$ och $r > 0$ så existerar $C > 0$ så att

$$|x|^r \leq C \cdot e^{|tx|} \forall x \in \mathbb{R}.$$

Därmed gäller också, för dessa värden på r och t,

$$E[|X|^r] \leq E[C \cdot E^{|tX|} < \infty.$$

(b): Taylorutveckling ger

$$e^{tX} = 1 + \sum_{n=1}^{\infty} \frac{t^n \cdot X^n}{n!}.$$

Givet att $|t| < h$ så existerar väntevärdet av ovan uttryck och ger

$$E[e^{tX}] = 1 + \sum_{n=1}^{\infty} \frac{t^n E[X^n]}{n!}.$$

Notera att högerledet är Taylorserien av VL. Därmed måste vi ha att

$$\Psi_X^{(n)}(0) = \frac{d^n}{(dt)^n} E[e^{tx}] = E[X^n].$$

$\square$

Exempel. Låt $Z \sim \mathrm{N}(0,1)$. För alla $t \in \mathbb{R}$ får vi att

$$\begin{aligned}
\Psi_Z(t) &= E[e^{tZ}] \\
&== \int_{-\infty}^{\infty} e^{tz} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2-2tz}{2}} dz \\
&= -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-t)^2}{2}+\frac{t^2}{2}} dz \\
&= e^{\frac{t^2}{2}}.
\end{aligned}$$

Sats 9.2. Låt X vara en stokastisk variabel och $a, b \in \mathbb{R}$. Då gäller

$$\Psi_{aX+b}(t) = e^{tb} \cdot \Psi_X(at),$$

givet att VL eller HL väldefinerat.

Bevis.

$$\begin{aligned}
\Psi_{aX+b}(t) &= E[e^{t(aX+b)}] \\
&= e^{tb} \cdot E[e^{(ta)X}] \\
&= e^{tb} \cdot \Psi_X(at).
\end{aligned}$$

$\square$

Exempel. Låt $Y \sim N(\mu, \sigma^2)$. Då är

$$z = \frac{Y-\mu}{\sigma} \sim N(0,1).$$

Därmed får vi

$$\begin{aligned}
\Psi_Y(t) &= \Psi_{\mu+\sigma Z}(t) \\
&= e^{\mu t}\Psi_Z(\sigma t) \\
&= e^{\mu t} \cdot e^{t^2 \frac{\sigma^2}{2}} \\
&= e^{\mu t + t^2\sigma^2/2}.
\end{aligned}$$

Sats 9.3. Låt $X_1, \ldots, X_n$ vara oberoende stokastiska variabler med MGF som existerar för $|t| < h$, för något $h > 0$. Låt $S_n = X_1 + \cdots + X_n$. Då gäller

$$\Psi_{S_n}(t) = \prod_{k=1}^{n} \Psi_{X_k}(t),$$

för $|t| < h$.

Sats 9.4. Om det existerar $h > 0$ så att $\Psi_X(t)$ och $\Psi_Y(t)$ existerar och är lika för $|t| < h$, då är X och Y likafördelade.

Dag

-1

Misc notes

10 Stuff

$N \sim \text{Po}(\lambda)$

$X|N = n$

11 Stuff 2

$\overline{X} \sim \text{Exp}(1/\alpha)$ (line of repairing device) $\overline{Y} \sim \text{Exp}(\frac{1}{n\alpha})$ (time of single aeorisoion?)

$$P_n = P(\overline{Y} \leq \overline{X})^n.$$

$$\begin{aligned} P(\overline{Y} \leq \overline{X}) &= P(\overline{Y} \leq \overline{X}|\overline{X} \in (0,\infty]) \\ &= \int_0^\infty P(\overline{Y} \leq \overline{X}|\overline{X} = x) f_X(x) dx \\ &= \int_0^\infty F_{\overline{Y}|\overline{X}=x}(x) f_{\overline{X}}(x) dx \\ &= \int_0^\infty P(\overline{Y} \leq x|\overline{X} = x) f_{\overline{X}}(x) dx \\ &= \int_0^\infty \left(\int_0^x F_{\overline{Y}|\overline{X}}(y) dy \right) f_{\overline{X}}(x) dx \\ &= \int_0^\infty \left(\int_0^x n\alpha e^{-n\alpha y} dy \right) a e^{-\alpha} dx. \end{aligned}$$

12 3.1a

$X|M = m \sim \text{Po}(m)$, $M \sim \text{Exp}(a)$. What is the distribution of X?

Think always when encountering this law of total probability.

$$P(X = k|M = m) = e^{-m}\frac{m^k}{k!}, k = 0, 1, \ldots$$
$$f_M(x) = \begin{cases} \frac{1}{a}e^{-\frac{x}{a}}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}.$$

Law of total probability: $B_1, \ldots, B_n$ disjoint events and $B_1 \cup \cdots \cup B_n = \Omega$. Discrete case:

$$P(A) = \sum_k P(A|B_k) \cdot P(B_k).$$

Continous case:

$$P(X = k) = \int_0^\infty P(X = k|M = m) f_M(m) dm.$$

Gamma distribution:

$$f(x,k,\theta) = \frac{x^{k-1}e^{-x/\theta}}{\theta^k\Gamma(k)}.$$

for $x \geq 0$, $k, \theta > 0$.

Gamma i $k!$.

It is needed when evaluating the integral.

13 Problem 5

N, X_1, X_2, independent random variables. $N \in \mathrm{Po}(\lambda)$. $X_k \in \mathrm{Be}(\frac{1}{2})$

$$Y_1 = \sum_{k=1}^{N} X_k$$
$$X_2 = NY_1.$$

Prove that Y_1 and Y_2 are independent. $y_1, y_2 \geq 0$ We want to prove

$$P(Y_1 = y, Y_2 = y_2) = P(Y_1 = y_1)P(Y_2 = y_2).$$

$P(Y_1 = y_1, N = y_1 + y_2)$.

You should condition in this case.

Condition on the random value.

14 Problem 19

$$f_{X,Y}(x,y) = \begin{cases} c, & 0 \leq x \leq 1, x^2 \leq y \leq x \\ 0, & \text{otherwise} \end{cases}.$$

$E[X|Y=y]$? $E[Y|X =>$

it ranges from $x^2 \leq y \leq x$.

kk