Föreläsning 13

den 18 maj 2021 13:04

Matematisk Statistik och Diskret Matematik, MVE051/MSG810, VT21

Timo Vilkas

Chalmers - Göteborgs Universitet

Föreläsning 13
18 maj 2021

Momentgenererande funktion

Som bekant ges momenten till en s.v. X som

$$\mathbb{E}(X^k) = \begin{cases} \sum_{x \in X(S)} x^k \cdot \mathbb{P}(X = x), & \text{om } X \text{ är diskret,} \\ \int_{-\infty}^{\infty} x^k \cdot f_X(x)\, dx, & \text{om } X \text{ är kontinuerlig.} \end{cases}$$

Den motsvarande momentgenererande funktionen (m.g.f.) är definerad som $m_X(t) = \mathbb{E}[e^{tX}] = \mathbb{E}\sum_{n=0}^{\infty} \frac{(tX)^n}{n!} \underset{\text{kons.}}{=} \sum_{n=0}^{\infty} \mathbb{E}(X^n) \cdot \frac{t^n}{n!} = g(t)$ för den exponentiella gen. funktionen g till talföljden $a_n = \mathbb{E}(X^n)$, $n \in \mathbb{N}_0$

Sats (jfr. föreläsning 2)

Låt X vara en slumpvariabel för vilken $\mathbb{E}[e^{tX}]$ existerar för t tillräckligt nära 0.

Då gäller $m_X^{(k)}(0) = \left.\frac{d^k m_X(t)}{dt^k}\right|_{t=0} = \mathbb{E}[X^k].$

Exempel

- I föreläsning 2 har vi hittat att $X \sim \text{Bin}(n, p)$ har den momentgenererande funktionen $m_X(t) = (q + pe^t)^n$, där $q = 1 - p$, $t \in \mathbb{R}$. Med

$$m_X'(t) = npe^t(q + pe^t)^{n-1} \quad \text{och}$$
$$m_X''(t) = n(n-1)p^2e^{2t}(q + pe^t)^{n-2} + npe^t(q + pe^t)^{n-1}$$

följer $\mathbb{E}\,X = m_X'(0) = n \cdot p$ och
$\mathbb{E}(X^2) = m_X''(0) = n(n-1)p^2 + np$ samt
$Var(X) = \mathbb{E}(X^2) - (\mathbb{E}\,X)^2 = n^2p^2 - np^2 + np - (np)^2 = np(1-p)$.

Exempel

- För $Z \sim \mathcal{N}(0,1)$ är m.g.f. $m_Z(t) = \exp\left(\frac{t^2}{2}\right)$:

$$m_Z(t) = \mathbb{E}(e^{tZ}) = \int_{-\infty}^{\infty} e^{tz} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\, dz$$

$$= e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z^2-2tz+t^2)}\, dz$$

variabelbyte
$y = z - t$
$dy = dz$

$$= e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy = e^{t^2/2}.$$

$= \mathbb{P}(Z \in \mathbb{R}) = 1$

Verktygslåda (1)

Sats (karakteriserande)

Låt X och Y vara två slumpvariabler med m.g.f. $m_X(t)$ resp. $m_Y(t)$. Om det gäller $m_X(t) = m_Y(t)$ för alla t i ett öppet intervall som innehåller $t = 0$, så har X och Y samma fördelning.

Bevis: utanför kursens räckvidd

Sats (linjär transformation)

Om X har m.g.f. $m_X(t)$ och $Y = aX + b$ för några $a, b \in \mathbb{R}$, så gäller $m_Y(t) = e^{bt}\, m_X(at)$.

Bevis: $m_Y(t) = \mathbb{E}\left(e^{t(aX+b)}\right) = \mathbb{E}\left(e^{tb} \cdot e^{(at)X}\right)$

$= e^{bt} \cdot \mathbb{E}\left(e^{(at)X}\right) = e^{bt} \cdot m_X(at).$ □

Exempel

För $X \sim \mathcal{N}(\mu, \sigma^2)$ är $Y = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$ (jfr. föreläsning 3).
Med $X = \sigma Y + \mu$, $m_Y(t) = \exp\left(\frac{t^2}{2}\right)$ och ovanstående sats följer

$$m_X(t) = \exp\left(\mu t + \frac{(\sigma t)^2}{2}\right).$$

Anmärkning: Man kan istället beräkna m.g.f. $m_X(t)$ till allmän $X \sim \mathcal{N}(\mu, \sigma^2)$ och använda båda satser ovan till ett enkelt bevis av $Y = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$ (vilket vi härledde på annan väg).

Verktygslåda (2)

Sats (summa blir produkt)

*Låt X och Y vara två slumpvariabler med m.g.f. $m_X(t)$ resp. $m_Y(t)$. Om X och Y är **oberoende**, så gäller*

$$m_{X+Y}(t) = m_X(t) \cdot m_Y(t).$$

"Bevis": Det enda som behöver verifieras är att $\exp(tX)$ och $\exp(tY)$ är oberoende ifall X och Y är det (lite tekniskt).
Sedan följer omedelbart:

$$\begin{aligned} m_{X+Y}(t) &= \mathbb{E}(e^{t(X+Y)}) = \mathbb{E}(e^{tX} \cdot e^{tY}) \\ &= \mathbb{E}(e^{tX}) \cdot \mathbb{E}(e^{tY}) = m_X(t) \cdot m_Y(t). \end{aligned}$$

Fördelningsfamiljer (1)

Proposition

Om $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ *och* $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ *är oberoende, så är* $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Bevis: Enligt satsen ovan är

$$m_{X_1+X_2}(t) = m_{X_1}(t) \cdot m_{X_2}(t) = \exp\left(\mu_1 t + \frac{(\sigma_1 t)^2}{2}\right) \cdot \exp\left(\mu_2 t + \frac{(\sigma_2 t)^2}{2}\right)$$
$$= \exp\left((\mu_1 + \mu_2)t + \frac{(\sigma_1^2+\sigma_2^2)t^2}{2}\right),$$

vilket är allt som behöver visas (enligt karakteriserande satsen ovan).

Fördelningsfamiljer (2)

Proposition

Om $X_1 \sim \text{Poiss}(\lambda_1)$ *och* $X_2 \sim \text{Poiss}(\lambda_2)$ *är oberoende, så är* $X_1 + X_2 \sim \text{Poiss}(\lambda_1 + \lambda_2)$.

Bevis: Poissonfördelningen med parameter λ har m.g.f. $e^{\lambda(e^t-1)}$ (jfr. föreläsning 4). Enligt satsen ovan är då

$$m_{X_1+X_2}(t) = m_{X_1}(t) \cdot m_{X_2}(t) = \exp\left(\lambda_1(e^t - 1)\right) \cdot \exp\left(\lambda_2(e^t - 1)\right)$$
$$= \exp\left((\lambda_1 + \lambda_2)(e^t - 1)\right),$$

vilket är allt som behöver visas.

Anmärkning: På exakt samma sätt, med m.g.f. $m(t) = (1 - \beta t)^{-\alpha}$ till $\Gamma(\alpha, \beta)$, kan man visa att $X_1 + X_2 \sim \Gamma(\alpha_1 + \alpha_2, \beta)$ för oberoende $X_1 \sim \Gamma(\alpha_1, \beta)$ och $X_2 \sim \Gamma(\alpha_2, \beta)$.

Väntevärdet är monotont

Lemma

För (antingen diskreta eller kontinuerliga) slumpvariabler X, Y med $X \leq Y$ gäller $\mathbb{E}\,X \leq \mathbb{E}\,Y$.

Bevis: $\mathbb{E}(Y - X) \geq 0$ följer direkt ur väntevärdets definition, med monotonin av summa/integral. Det räcker faktiskt att $\mathbb{P}(X \leq Y) = 1$ (vilket är lite svagare än $X \leq Y$).

Exempel

$X \sim \mathrm{Geom}(p)$, $Y \sim \mathrm{NB}(r, p)$ för $r \in \mathbb{N}$, $p \in (0,1) \Rightarrow \mathbb{E}\,X \leq \mathbb{E}\,Y$.
Vi har redan kunskap om väntevärden och vet att denna olikhet gäller. Dock följer det här utan att ens beräkna ett väntevärde. Om X räknar antalet försök till första succée och Y antalet försök till succée nummer $r \geq 1$ i **samma** följd av oberoende $\mathrm{Ber}(p)$-experiment, så gäller $X \leq Y$ och lemmat kan användas.

Chebyshevs olikhet

Sats

Låt $\varepsilon > 0$ och X vara en s.v. med väntevärde $\mathbb{E}\,X = \mu$ samt varians $Var(X) = \sigma^2$. Då gäller

$$\mathbb{P}(|X - \mu| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}.$$

Bevis: Låt $Y := \begin{cases} \varepsilon^2, & \text{om } |X-\mu| \geq \varepsilon, \\ 0, & \text{annars.} \end{cases}$ $\Rightarrow |X-\mu|^2 \geq Y$

$\mathbb{E}\,Y = 0 \cdot \mathbb{P}(|X-\mu| < \varepsilon) + \varepsilon^2 \cdot \mathbb{P}(|X-\mu| \geq \varepsilon) \Rightarrow \mathbb{P}(|X-\mu| \geq \varepsilon) = \frac{\mathbb{E}Y}{\varepsilon^2} \leq \frac{var(X)}{\varepsilon^2}$

$\mathbb{E}Y \overset{\text{Lemma}}{\leq} \mathbb{E}(|X-\mu|^2) = \mathbb{E}((X-\mathbb{E}X)^2) = var(X)$ □.

Anmärkning: Man kan välja $\varepsilon = k\sigma$ och får då $\mathbb{P}(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$ för godtyckligt $k > 0$ istället.

Stora talens lag (STL)

Sats

Låt $X_1, X_2, \ldots, X_n$ vara oberoende slumpvariabler som har samma väntevärde $\mathbb{E}\, X_i = \mu$ och varians $Var(X_i) = \sigma^2$, där $1 \leq i \leq n$ (t.ex. ett stickprov ur samma fördelning).
För $S_n = \sum_{i=1}^{n} X_i$ och godtyckligt $\varepsilon > 0$ gäller då

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \to 0 \quad \textit{med } n \to \infty.$$

Bevis: Låt $Y := \frac{S_n}{n}$. Då gäller $\mathbb{E}\, Y = \mu$ och (X_i oberoende) $Var(Y) = \frac{1}{n^2}\, Var(S_n) = \frac{\sigma^2}{n}$. Med Chebyshevs olikhet följer

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{\varepsilon^2 n} \to 0 \quad \text{med } n \to \infty.$$

Anmärkning: Beviset av STL ger oss t.o.m. en konvergenshastighet (svanssannolikheten avtar minst $\sim \frac{1}{n}$), vilken i många situationer dock är långt ifrån den verkliga hastigheten (som ofta är $\sim \frac{1}{n^2}$).

Exempel

Låt $(X_n)_{\in \mathbb{N}}$ vara en följd av oberoende tärningskast (symmetrisk tärning med 6 sidor). Låt $A = \{3, 4, 5, 6\}$ vara händelsen att tärningen visar minst 3 ögon. Visa att den **relativa frekvensen** *(eng. relative frequency)*, dvs. antalet gånger A inträdde bland de första n försök, delat på n, ligger nära $p := \mathbb{P}(A)$ för stora n.

Låt $Y_n := \begin{cases} 1, & \text{om } X_n \in A, \\ 0, & \text{annars.} \end{cases} = \mathbb{1}_A(X_n)$ och $S_n = \sum_{i=1}^{n} Y_i$ för $n \in \mathbb{N}$.

Då är $\frac{S_n}{n}$ den relativa frekvensen och $\mathbb{E}\left(\frac{S_n}{n}\right) = \mathbb{E}\, Y_1 = p = \frac{2}{3}$, samt $Var(Y_1) = p(1-p) = \frac{2}{9}$. Enligt STL gäller för godtyckligt $\varepsilon > 0$:

$$\mathbb{P}\left(\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right) \leq \frac{2}{9\varepsilon^2 \cdot n} \longrightarrow 0 \quad \text{när } n \to \infty.$$