



Kapitel 3

Diskreta slumpvariabler och deras sannolikhetsfördelningar



Diskreta slumpvariabler

En slumpvariabel tilldelar tal till samtliga utfall i ett slutförsök.

Vi kan skapa en *sannolikhetsfunktion* för en slumpvariabel X via

$$p(x) = \Pr(X = x) = \Pr(A)$$

där A består av alla utfall i utfallsrummet som X värderar till x .



Väntevärden för diskreta slumpvariabler

Vi är ofta intresserade av det värde en slumpvariabel *förväntas* anta, dvs det värde den antar i genomsnitt då slumpförsöket utförs vid upprepade tillfällen.

$$\mu = E(X) = \sum xp(x)$$

där summationen görs över alla x som X kan anta.

Låt $u(X)$ vara en funktion av X . Väntevärdet för $u(X)$ blir

$$E(u(X)) = \sum u(x)p(x)$$

där summationen åter görs över alla x som X kan anta.



Väntevärden för diskreta slumpvariabler

$E(X)$ anger det värde X antar i *genomsnitt*. Då slumpförsöket upprepas kommer de värden X antar att variera. Som ett mått på graden av variation används *variansen* av X .

$$\sigma^2 = V(X) = E[X - E(X)]^2 = E(X - \mu)^2 = \sum (x - \mu)^2 p(x)$$

Variansen för X är väntevärdet för $u(X)=(X-\mu)^2$, och mäter alltså *det förväntade kvadratavståndet mellan X och sitt väntevärde*.

För att få ett mått på spridning som lämpar sig för tolkning tas kvadratroten ur variansen. Detta ger *standardavvikelsen* σ som därmed är ett mått på *det förväntade avståndet mellan X och μ* .



Beräkningsformler för $V(X)$

Definitionsformeln för $V(X)$ är ofta besvärlig att arbeta med och som väl är finns det beräkningsformler. Vi vet sedan tidigare att

$$V(X) = E(X^2) - E^2(X) = E(X^2) - \mu^2$$

Många gånger är det dock mer lämpligt att istället använda följande samband.

$$V(X) = E[X(X-1)] + E(X) - E^2(X)$$



Vanligt förekommande diskreta slumpvariabler

I många situationer är vi hos individerna i en population enbart intresserade av huruvida dom har en viss egenskap eller inte.

Betraktar vi det som en urnmodell får vi en urna med bollar i två färger (exempelvis vita och svarta bollar).

När vi nu på något sätt drar bollar ur urnan och sedan på något sätt värderar de resulterande utfallen uppstår intressanta slumpvariabler.



Binomialfördelningen

En urna innehåller enbart vita och svarta bollar. Andelen vita bollar i urnan är p . Vi drar *med* återläggning n bollar ur urnan och studerar slumpvariabeln

$X = \text{Antal vita bollar i urvalet}$

X sägs vara *binomialfördelad* med parametrar n och p ; $Bi(n,p)$.

Sannolikhetsfunktion, väntevärde och varians ges av

$$p(x) = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, \dots, n$$
$$\begin{aligned} \mu &= E(X) = np \\ \sigma^2 &= V(X) = npq \end{aligned}$$



Geometrisk fördelning

En urna innehåller enbart vita och svarta bollar. Andelen vita bollar i urnan är p . Vi drar *med* återläggning bollar ur urnan tills vi för första gången får en vit boll och studerar slumpvariabeln

$X = \text{Antal dragna bollar}$

X sägs vara *geometriskt fördelad* med parameter p ; $Ge(p)$.

Sannolikhetsfunktion, väntevärde och varians ges av

$$p(x) = pq^{x-1}, \quad x = 1, 2, \dots$$
$$\mu = E(X) = \frac{1}{p}$$
$$\sigma^2 = V(X) = \frac{q}{p^2}$$



Matematiskt verktyg: Derivering i samband med serier

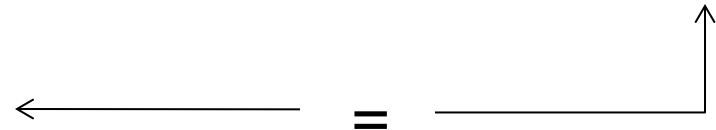
Betrakta en geometrisk serie som en funktion av q

$$f(q) = \sum_{k=0}^{\infty} q^k = \frac{1}{1-q}, \text{ d\u00e4r } |q| < 1$$

Vi deriverar v\u00e4nsterled och h\u00f6gerled var f\u00f6r sig

$$f'(q) =$$

$$f'(q) = \frac{d}{dq} \left(\frac{1}{1-q} \right)$$





Geometrisk fördelning. Bevis för väntevärdet

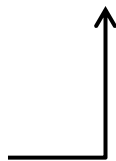
Låt X vara $Ge(p)$. Då följer att

Definition av väntevärde.

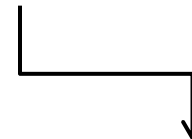


$$E(X) = \sum_{x=1}^{\infty} x p q^{x-1}$$

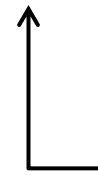
Eftersom p enbart är en konstant kan den flyttas utanför summan.



Vi förenklar uttrycket i nämnaren via sambandet $p = 1 - q$.



Nu använder vi sambandet från föregående bild.





Uppgift 3.77

Låt Y vara $Ge(p)$. Bestäm sannolikheten att Y antar ett udda värde, dvs $y=1,3,5,\dots$

För att Y ska anta ett *udda* värde måste det ha kommit ett *jämnt* antal svarta bollar före den vita bollen. Sannolikheterna blir

$$p, pq^2, pq^4, pq^6, \dots$$

vilket leder till att

$$\Pr(Y = \text{“Udda”})$$



Negativ binomialfördelning

En urna innehåller enbart vita och svarta bollar. Andelen vita bollar i urnan är p . Vi drar *med* återläggning bollar ur urnan tills vi för r :te gången får en vit boll och studerar slumpvariabeln

$X = \text{Antal dragna bollar}$

X sägs följa en *negativ binomialfördelning* med parametrar r och p ; $\text{NegBin}(r, p)$.

$$p(x) = \binom{x-1}{r-1} p^r q^{x-r}, \quad x = r, r+1, \dots$$
$$\mu = E(X) = \frac{r}{p}$$
$$\sigma^2 = V(X) = \frac{rq}{p^2}$$



Hypergeometrisk fördelning

En urna innehåller enbart vita och svarta bollar. Antalet bollar i urnan är N och antalet vita bollar i urnan är r . Vi drar *utan återläggning* n bollar ur urnan och studerar slumpvariabeln

$X =$ *Antal vita bollar i urvalet*

X sägs vara *hypergeometriskt fördelad* med parametrar n , r och N ; $Hyp(n, r, N)$.

$$p(x) = \frac{\binom{r}{x} \binom{N-r}{n-x}}{\binom{N}{n}}, \quad x = 0, 1, \dots, n$$

$$\mu = E(X) = \frac{nr}{N}$$

$$\sigma^2 = V(X) = n \left(\frac{r}{N} \right) \left(\frac{N-r}{N} \right) \left(\frac{N-n}{N-1} \right)$$



Poissonprocesser

Antag att händelser inträffar på ett intervall enligt följande axiom

1. Händelser inträffar med en viss intensitet λ .
2. Antal händelser på disjunkta intervall är oberoende.
3. På ett tillräckligt kort intervall är sannolikheten för exakt en händelse approximativt proportionell mot intensiteten λ och intervalllets längd h , dvs sannolikheten är approximativt λh . På ett sådant kort intervall gäller vidare att sannolikheten för två eller fler händelser approximativt är noll. Således gäller att sannolikheten för att ingen händelse inträffar på ett sådant intervall approximativt är $1 - \lambda h$.

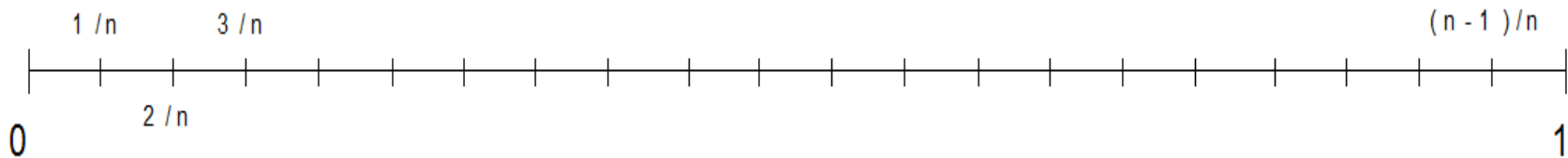
Vilken sannolikhetsfördelning har slumpvariabeln

$X = \text{Antal händelser på ett enhetsintervall}$



Poissonfördelningen: Samband med Binomialfördelningen

Vi delar in enhetsintervallet i n lika korta delintervall.



Enligt axiomen följer då att X approximativt är $Bi(n, \lambda/n)$, dvs

$$p(x) \approx \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

Denna approximation bör bli bättre och bättre få n ökar. Den exakta sannolikhetsfunktionen fås som gränsvärde då $n \rightarrow \infty$.



Matematiskt verktyg: Gränsvärden i samband med de fyra räknesätten

Betrakta två funktioner f och g med följande gränsvärden.

$$\lim_{x \rightarrow \infty} f(x) = A$$

$$\lim_{x \rightarrow \infty} g(x) = B$$

Om vi nu bildar en ny funktion av f och g med något av de fyra räknesätten visar det sig att motsvarande nybildning gäller för gränsvärdet.

$$\lim_{x \rightarrow \infty} [f(x) + g(x)] = A + B$$

$$\lim_{x \rightarrow \infty} [f(x) - g(x)] = A - B$$

$$\lim_{x \rightarrow \infty} [f(x) g(x)] = AB$$

$$\lim_{x \rightarrow \infty} \left[\frac{f(x)}{g(x)} \right] = \frac{A}{B}$$



Poissonfördelningen: Sannolikhetsfunktionen

Den sannolikhetsfunktion som söks är gränsvärdet

$$p(x) = \lim_{n \rightarrow \infty} \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

För att kunna använda räkneregler för gränsvärden stuvlar vi om i uttrycket och får produkten

$$\lim_{n \rightarrow \infty} \frac{\lambda^x}{x!} \cdot \frac{n(n-1) \cdots (n-x+1)}{n^x} \cdot \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x}$$



Poissonfördelningen: Sannolikhetsfunktionen

Termen $\lambda^x/x!$ påverkas inte av $n \rightarrow \infty$. Vidare gäller att

$$\lim_{n \rightarrow \infty} \frac{n(n-1) \cdots (n-x+1)}{n^x}$$

Sedan har vi det i matematiken välkända gränsvärdet

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x \quad \text{som ger att} \quad \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}$$

I den sista termen är det bara λ/n som påverkas så

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^{-x} = (1 - 0)^{-x} = 1$$



Poissonfördelningen

Antag att händelser inträffar på ett intervall enligt tidigare angivna axiom. Betrakta slumpvariabeln

$X =$ Antal händelser på ett enhetsintervall

Sannolikhetsfunktion, väntevärde och varians ges av

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \dots \quad \begin{array}{l} \mu = E(X) = \lambda \\ \sigma^2 = V(X) = \lambda \end{array}$$



Moment

$E(X)$ kallas för det första *momentet* för X . Det k .te momentet för X definieras via.

$$\mu'_k = E(X^k) = \sum x^k p(x)$$

Då man studerar moment kring väntevärdet fås sk *centralmoment* för X . Det k .te centralmomentet för X definieras via.

$$\mu_k = E(X - \mu)^k = \sum (x - \mu)^k p(x)$$

Vid viss bevisföring har vi även nytta av sk *faktorialmoment*. Det k .te faktorialmomentet för X definieras via.

$$\mu_{[k]} = E[X(X-1)\cdots(X-k+1)]$$



Momenten ger sannolikhetsfördelningen

$E(X)$ och $V(X)$ används ofta för att *sammanfatta* en sannolikhetsfördelning.

De två första momenten, dvs $E(X)$ och $E(X^2)$, räcker emellertid *inte* för att *fullständigt* beskriva denna sannolikhetsfördelning.

Dock gäller (i princip) att *samtliga* moment, dvs $E(X)$, $E(X^2)$, $E(X^3)$, \dots , är tillräckligt för att *fullständigt* beskriva sannolikhetsfördelningen för X .

Alltså; om samtliga moment för två slumpvariabler X och Y existerar och är samma, dvs $E(X)=E(Y)$, $E(X^2)=E(Y^2)$, $E(X^3)=E(Y^3)$, \dots , så gäller (i princip) att X och Y följer samma sannolikhetsfördelning.

Om vi kan finna en funktion som är uppbyggd av momenten för X kommer denna funktion att entydigt bestämma sannolikhetsfördelningen för X .



Matematiskt verktyg: Taylorutveckling

En funktion f som är oändligt deriverbar i punkten $x=a$ kan beskrivas med en *polynomserie* via en sk *Taylorutveckling* av f .

$$f(x) = f(a) + \frac{f'(a)(x-a)}{1!} + \frac{f''(a)(x-a)^2}{2!} + \dots = \sum_{k=0}^{\infty} \frac{f^{(k)}(a) \cdot (x-a)^k}{k!}$$

Specialfallet $a=0$ kallas för *Maclaurinutveckling* av f .

$$f(x) = f(0) + \frac{f'(0)x}{1!} + \frac{f''(0)x^2}{2!} + \dots = \sum_{k=0}^{\infty} \frac{f^{(k)}(0) \cdot x^k}{k!}$$

För $f(x)=e^x$ gäller att $f^{(n)}(0)=e^0=1$ för alla n . *Maclaurinutvecklingen* blir

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$



Momentgenererande funktioner

Låt X vara en slumpvariabel där samtliga moment existerar. Den *momentgenererande* funktionen för X definieras via

$$m(t)$$

Den momentgenererande funktionen för X är alltså en funktion av t och existerar endast då den är ändlig i en omgivning av origo.



Momentgenererande funktioner

Varför kallas $m(t)$ för en momentgenererande funktion för X ? Låt oss börja med att göra en Maclaurinutveckling av e^{tx} .

$$e^{tx} = 1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \dots$$

Under förutsättning av samtliga moment för X existerar gäller att

$$m(t) = E(e^{tX})$$



Momentgenererande funktioner

Den momentgenererande funktionen $m(t)$ är alltså uppbyggd av momenten för X . Vi finner dessa moment via upprepad derivering av $m(t)$. Eftersom

$$m(t) = 1 + tE(X) + \frac{t^2}{2!} \cdot E(X^2) + \frac{t^3}{3!} \cdot E(X^3) + \dots$$

följer att förstaderivatans blir

$$m'(t) = E(X) + \frac{2t}{2!} \cdot E(X^2) + \frac{3t^2}{3!} \cdot E(X^3) + \frac{4t^3}{4!} \cdot E(X^4) \dots$$

Eftersom t ingår i alla termer utom den första följer att $m'(0) = E(X)$.



Momentgenererande funktioner

Förstaderivatans $m'(t)$ är en "förskjutning" av $m(t)$. På motsvarande sätt kommer andraderivatans $m''(t)$ att vara en förskjutning av $m'(t)$, dvs

$$m''(t) = E(X^2) + t \cdot E(X^3) + \frac{t^2}{2!} \cdot E(X^4) + \frac{t^3}{3!} \cdot E(X^5) \dots$$

Eftersom t ingår i alla termer utom den första följer att $m''(0) = E(X^2)$.

Fortsätter vi på samma sätt följer att

$$m^{(k)}(t) = E(X^k) + t \cdot E(X^{k+1}) + \frac{t^2}{2!} \cdot E(X^{k+2}) + \frac{t^3}{3!} \cdot E(X^{k+3}) + \dots = \sum_{j=k}^{\infty} \frac{t^j}{j!} \cdot E(X^j)$$

Eftersom t ingår i alla termer utom den första följer att $m^{(k)}(0) = E(X^k)$.



Uppgift 3.145

Låt Y vara $Bi(n,p)$, dvs sannolikhetsfunktionen för Y ges av

$$p(y) = \binom{n}{y} p^y q^{n-y}, \quad y = 0, 1, 2, \dots, n$$

Den momentgenererande funktionen för Y blir därmed

$$m(t) = \sum_{y=0}^n e^{ty} \binom{n}{y} p^y q^{n-y}$$

Hur skriver vi om detta på sluten form?

Vi slår ihop termerna e^{ty} och p^y .

Vi utnyttjar att summan är en binomialutveckling av typen $(a+b)^n$.



Matematiskt verktyg: Derivatans av en produkt och kedjeregeln

Låt f och g vara två deriverbara funktioner. Då gäller att

$$(fg)'(x) = [f(x)g(x)]' = f'(x)g(x) + f(x)g'(x)$$

Betrakta den sammansatta funktionen $h(x)=g(f(x))$.

”Kedjeregeln” ger att

$$[g(f(x))]' = g'(f(x))f'(x)$$

Låter vi $y=f(x)$ får vi med den sk *Leibniz-notationen* att

$$\frac{dg}{dx} = \frac{dg}{dy} \cdot \frac{dy}{dx}$$



Uppgift 3.146

För att bestämma $E(Y)$ och $V(Y)$ beräknas första- och andraderivatn av $m(t)$. Vi får först att

$$m'(t) = n (pe^t + q)^{n-1} \cdot pe^t$$

vilket leder till att

$$E(Y) = m'(0) = np (pe^0 + q)^{n-1} \cdot e^0$$

eftersom $e^0=1$



Uppgift 3.146

Andraderivatan av $m(t)$ blir (efter lite bök)

$$m''(t) = n(n-1)p^2(pe^t + q)^{n-2} \cdot e^{2t} + np(pe^t + q)^{n-1} \cdot e^t$$

vilket leder till att

$$E(Y^2) = m''(0) = n(n-1)p^2 + np$$

Beräkningsformeln $V(Y) = E(Y^2) - E^2(Y)$ ger därmed att

$$V(Y) = n(n-1)p^2 + np - n^2p^2 = np - np^2 = \mathbf{npq}$$



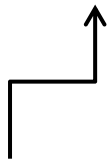
Uppgift 3.147

Låt Y vara $Ge(p)$, dvs sannolikhetsfunktionen för Y ges av

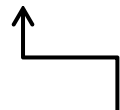
$$p(y) = pq^{y-1}, \quad y = 1, 2, \dots$$

Den momentgenererande funktionen för Y blir därmed

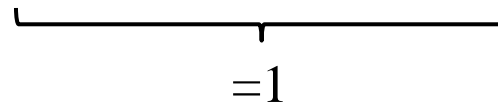
$$m(t) = \sum_{y=1}^{\infty} e^{ty} pq^{y-1}$$



Hur skriver vi om detta på sluten form?



Summanden uppvisar stora likheter med slh-funktionen för en $Ge(1-qe^t)$ varför vi modifierar på vanligt sätt.





Matematiskt verktyg: Derivatans av en kvot

Låt f och g vara två deriverbara funktioner (med $g(x) \neq 0$). Då gäller att

$$\left(\frac{f}{g}\right)'(x) = \left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$$



Uppgift 3.148

För att bestämma $E(Y)$ och $V(Y)$ beräknas första- och andraderivatn av $m(t)$. Vi får först att

$$m'(t) = \frac{pe^t(1 - qe^t) + pe^tqe^t}{(1 - qe^t)^2}$$

vilket leder till att

$$E(Y) = m'(0) = \frac{p}{(1 - q)^2}$$

eftersom $e^0=1$



Uppgift 3.148

Andraderivatatan av $m(t)$ blir (efter lite bök)

$$m''(t) = \frac{pe^t(1 - qe^t)^2 + 2pe^tqe^t(1 - qe^t)}{(1 - qe^t)^4}$$

vilket leder till att

$$E(Y^2) = m''(0) = \frac{p(1 + q)}{(1 - q)^3}$$

Beräkningsformeln $V(Y) = E(Y^2) - E^2(Y)$ ger därmed att

$$V(Y) = \frac{1 + q}{p^2} - \frac{1}{p^2}$$



Den momentgenererande funktionen ger sannolikhetsfördelningen

Via den momentgenererande funktionen kan man oftast någorlunda "enkelt" finna momenten för en slumpvariabel X , vilket bl a innebär att $E(X)$ och $V(X)$ kan bestämmas.

Den viktigaste egenskapen hos momentgenererande funktioner är emellertid att den *fullständigt* och *entydigt* bestämmer sannolikhetsfördelningen för X .

Alltså; om två slumpvariabler X och Y har momentgenererande funktioner som är identiska, dvs om det finns ett b sådant att

$$m_X(t) = m_Y(t), \quad |t| < b, \quad b > 0$$

gäller att X och Y följer samma sannolikhetsfördelning.



Uppgift 3.149

Vilken sannolikhetsfördelning följer Y om den momentgenererande funktionen för Y är

$$m(t) = (0.6e^t + 0.4)^3$$

Eftersom den momentgenererande funktionen för binomialfördelningen är

$$m(t) = (pe^t + q)^n$$

följer av entydigheten hos momentgenererande funktioner att Y är binomialfördelad med $n=3$ och $p=0.6$, dvs $Bi(3,0.6)$.



Momentgenererande funktioner för linjärfunktioner

Vi har tidigare sett hur man utifrån $E(Y)$ och $V(Y)$ kan bestämma $E(aY+b)$ och $V(aY+b)$ där a och b är konstanter.

Om vi på motsvarande sätt utifrån $m_Y(t)$ kan bestämma den momentgenererande funktionen för $aY+b$ har vi entydigt bestämt sannolikhetsfördelningen för $aY+b$.



Momentgenererande funktioner för linjärfunktioner (Uppgift 3.158)

Låt Y ha momentgenererande funktion $m(t)$. Betrakta linjärfunktionen $W=aY+b$. Då gäller att

$$m_W(t)$$



Uppgift 3.160

Låt Y vara $Bi(n,p)$. Bestäm den momentgenererande funktionen för $Y^* = n - Y$.

Y^* är en linjärfunktion av Y , dvs $Y^* = (-1) \cdot Y + n$, dvs $a = -1$, $b = n$.

Eftersom $m_Y(t) = (pe^t + q)^n$ följer att

$$m_{Y^*}(t)$$

Entydigheten ger att Y^* är $Bi(n,q)$ där $q = 1 - p$.



Sannolikhetsgenererande funktioner

De mest använda diskreta sannolikhetsfördelningarna kan enbart anta icke-negativa heltalsvärden. För dessa sannolikhetsfördelningar finns ytterligare en användbar "genererande" funktion.

Definition. Låt X vara en *diskret* slumpvariabel som enbart kan anta icke-negativa *heltalsvärden* och där samtliga *faktorialmoment* existerar. Den *sannolikhetsgenererande* funktionen för X definieras via

$$P(t)$$

Anledningen till att $P(t)$ kallas för sannolikhetsgenererande funktion är att vi (i likhet med mgf) via derivering kan finna sannolikheterna för X . Det gäller att

$$p(k) = \frac{P^{(k)}(0)}{k!}$$



Sannolikhetsgenererande funktioner

Vad har vi då för nytta av $P(t)$? Upprepad derivering ger svaret.

$$P(t) = p(0) + p(1)t + p(2)t^2 + \dots = \sum_{x=0}^{\infty} t^x p(x)$$

Deriveringsregeln för polynom ger att förstaderivatan blir

$$P'(t) = p(1) + 2p(2)t + 3p(3)t^2 + \dots$$

Betraktar vi nu denna derivata för $t=1$ följer att

$$P'(1) = p(1) + 2p(2) + 3p(3) + \dots$$



Vidare derivering ger att

$$P''(t)$$

och betraktar vi nu andraderivatan för $t=1$ följer att

$$P''(1) = \sum_{x=2}^{\infty} x(x-1)p(x)$$

$P''(1)$ är alltså det andra *faktorialmomentet* för X .



Sannolikhetsgenererande funktioner

Om vi fortsätter vår derivering av $P(t)$ följer att

$$P^{(k)}(t) = \sum_{x=k}^{\infty} x(x-1)\cdots(x-k+1)t^{x-k}p(x)$$

och således blir

$$P^{(k)}(1) = \sum_{x=k}^{\infty} x(x-1)\cdots(x-k+1)p(x)$$

Funktionen $P(t)$ kan alltså sägas generera faktorialmoment för X och faktum är att en annan benämning av $P(t)$ är *faktorialmomentgenererande funktion*.



Uppgift 3.165

Låt Y vara $Po(\lambda)$. Den sannolikhetsgenererande funktionen för Y blir då

$$P(t) = \sum_{y=0}^{\infty} t^y \frac{\lambda^y e^{-\lambda}}{y!}$$

$p(y)$ $\lambda^y \cdot t^y = (\lambda t)^y$
↓ ↓

↑

Eftersom termerna i summan är sannolikheter för $Po(\lambda t)$ så blir summan 1.



Chebyshevs olikhet

Detta avsnitt läses på egen hand. Vi återkommer till temat i Kapitel 4.