Sannolikhetsgenererande funktion

Den sannolikhetsgenererande funktionen är en transform av sannolikhetsfördelningen som inom matematik brukar kallas z-transform. Den definieras som

$$g_X(t) = E(t^X)$$

som ses som en funktion av t. Den är mest användbar för diskreta stokastiska variabler som kan anta värdena $0, 1, 2, \cdots$ och blir då

$$g_X(t) = \sum_{k=1}^{\infty} p_X(k) t^k.$$

För diskreta variabler av ovanstående typ är alltså den sannolikhetsgenererande funktionen en potensserie eller t o m ett polynom om X bara antar ett ändligt antal värden. Koefficienter för t^k är alltså $p_X(k) = P(X = k)$.

Exempelvis svarar mot "tärningskastfördelningen" som har $p_X(k) = 1/6,\ k = 1, 2, \cdots, 6$ funktionen $g_X(t) = (t + t^2 + \cdots + t^6)/6$.

För Poissonfördelningen $\mathrm{Po}(m)$ med sannolikhetsfunktionen $p_X(k) = m^k/k!\, e^{-m}$ får man den sannolikhetsgenererande funktionen till

$$g_X(t) = \sum_{k=0}^{\infty} \frac{m^k}{k!} e^{-m} t^k = e^{-m} \sum_{k=0}^{\infty} \frac{(mt)^k}{k!} = e^{-m} e^{mt} = e^{-m(1-t)}.$$

Eftersom $p_X(k) \geq 0$ och $\sum_0^\infty p_X(k) = 1$ konvergerar potensserien åtminstone om $|t| \leq 1$. Potensserier och polynom bestäms av sina koefficienter och det betyder att om en sannolikhetsgenererande funktion är given i potensserieform kan man direkt avläsa sannolikheterna eftersom koefficienten framför t^k är just $P(X = k)$. Det finns en en-entydig motsvarighet mellan diskreta fördelningar och sannolikhetsgenererande funktioner.

Användbarheten hos den sannolikhetsgenererande funktionen framgår då man gör observationen att om X och Y är oberoende så är

$$g_{X+Y}(t) = E(t^{X+Y}) = E(t^X t^Y) = (\text{oberoendet}) = E(t^X)E(t^Y) = g_X(t) g_Y(t).$$

Detta gör att man lätt kan få fram sannolikhetsfördelningen för summan av oberoende diskreta stokastiska variabler genom att 1) beräkna deras sannolikhetsgenererande funktioner, 2) multiplicera dessa, 3) skriva produkten som en potensserie samt 4) identifiera koefficienten framför t^k som $P(X + Y = k)$. Kan man direkt identifiera produkten som en viss fördelnings sannolikhetsfunktion kan man direkt se vilken fördelning summan har. Detta är en mycket enklare metod än att falta fördelningarna. Mot den komplicerade operationen "faltning" svarar den enkla operationen "multiplikation" av transformerna.

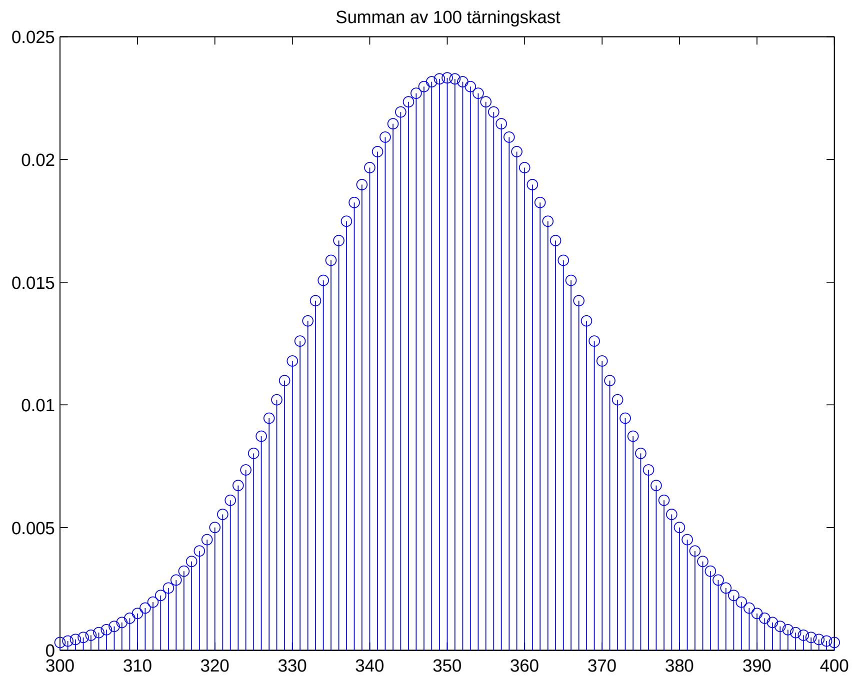
Tillämpat på Poissonfördelningen ger detta att om X är $\mathrm{Po}(m_1)$ och Y är $\mathrm{Po}(m_2)$ och oberoende så får vi $g_X(t) = e^{-m_1(1-t)}$ och $g_Y(t) = e^{-m_2(1-t)}$. Vidare är då

$$g_{X+Y}(t) = g_X(t) g_Y(t) = e^{-m_1(1-t)} e^{-m_2(1-t)} = e^{-(m_1+m_2)(1-t)}$$

som kan serieutvecklas för att erhålla $P(X + Y = k) = (m_1 + m_2)^k/k!\, e^{-(m_1+m_2)}$, dvs vi ser att $X + Y$ är $\mathrm{Po}(m_1 + m_2)$. Egentligen behöver vi inte ens serieutveckla utan ser direkt att

$X+Y$ har en sannolikhetsgenererande funktion som vi kan identifiera med $\text{Po}(m_1+m_2)$-fördelningens. Detta utnyttjar att serieutvecklingen är entydig.

Matlab har funktionen `conv` som utför multiplikation av polynom. Om vi lagrar sannolikheterna för ett tärningskast i vektorn `x=[0 1/6 1/6 1/6 1/6 1/6 1/6]` så ger `conv(x,x)` som resultat en 13-dimensionell vektor som innehåller sannolikheterna för att summan av två tärningskast skall bli $0, 1, 2, 3, \cdots, 12$. På motsvarande sätt ger Matlab-koden




```
last=x ;
for i=1:99,
last=conv(x,last);
end ;
```

i vektorn `last` sannolikhetsfördelningen för summan av 100 tärningskast. Den centrala delen av denna fördelning framgår av figuren.