



GÖTEBORGS
UNIVERSITET

MVE051/MSG810 - Statistik och diskret matematik (LP4 2022-2023)

Chalmers och Göteborgs universitet

Föreläsning 4 – Fler användbara fördelningar

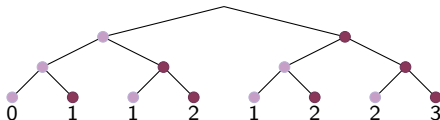
March 22, 2023

Dagens föreläsning

- Diskreta stokastiska variabler
 - Bernoullifördelning
 - Binomialfördelning
 - Geometrisk fördelning
 - Negativ binomialfördelning
 - Hypergeometrisk fördelning
 - Poissonfördelning
- Kontinuerliga stokastiska variabler
 - Likformig fördelning
 - Normalfördelning
 - Gamma-fördelning
 - Exponentialfördelning
 - χ^2 -fördelning

Binomialfördelning

Anta att X är antalet lyckade försök av totalt n . Anta att försöken är identiska, oberoende av varandra, och lyckas med sannolikhet p .

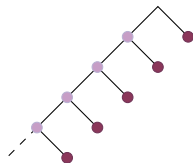


Då är X binomialfördelad med parametrar n och p (vi skriver $X \sim \text{bin}(p)$) och motsvarande sannolikhetsfunktion är

$$f(k) = P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k \in \{0, 1, \dots, n\}.$$

Geometrisk fördelning

Anta att X beskriver antalet gånger vi behöver upprepa ett försök tills det lyckas. Anta att försöken är identiska, oberoende av varandra, och lyckas med sannolikhet p .

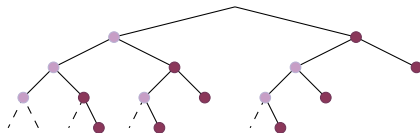


Då har X en geometrisk fördelning med parameter p (vi skriver $X \sim \text{geom}(p)$), och motsvarande sannolikhetsfunktion ges av

$$f(k) = P(X = k) = (1 - p)^{k-1}p, \quad k \geq 1.$$

Negativ binomialfördelning

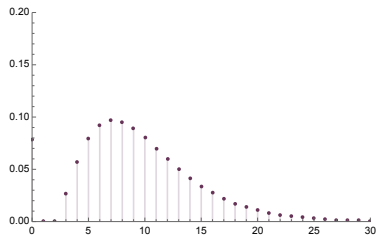
Anta att X beskriver antalet gånger vi behöver upprepa ett försök tills r försök har lyckats. Anta att försöken är identiska, oberoende av varandra, och lyckas med sannolikhet p .



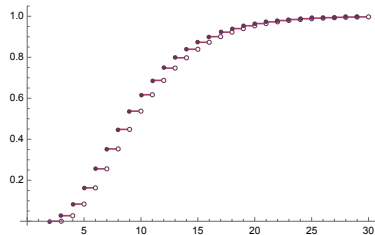
Då har X en negativ binomialfördelning med parametrar r och p , (vi skriver $X \sim \text{NB}(r, p)$) och motsvarande sannolikhetsfunktion ges av

$$f(k) = P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k \geq r$$

Negativ binomialfördelning



$$f(k) = P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$



$$F(k) = P(X \leq k)$$

Negativ binomialfördelning

Om $X \sim \text{NB}(r, p)$ så kan vi skriva

$$X = X_1 + \cdots + X_r,$$

där $X_1, \dots, X_r$ är oberoende stokastiska variable med fördelning $\text{geom}(p)$.

Alltså är

$$\mathbb{E}[X] = \mathbb{E}[X_1 + \cdots + X_r] = \underbrace{\frac{1}{p} + \cdots + \frac{1}{p}}_{r \text{ st.}} = \frac{r}{p}$$

$$\text{Var}(X) = \text{Var}(X_1 + \cdots + X_r) = \text{Var}(X_1) + \cdots + \text{Var}(X_r) = \frac{r(1-p)}{p^2}$$

och

$$\mathbb{E}[e^{tX}] = \mathbb{E}[e^{t(X_1 + \cdots + X_r)}] = \mathbb{E}[e^{tX_1}] \cdots \mathbb{E}[e^{tX_r}] = \left(\frac{pe^t}{1 - (1-p)e^t} \right)^r$$

Negativ binomialfördelning

Exempel

Antag att vi kastar en tärning.

1. Vad är sannolikheten att vi får en sexa först på det tredje kastet?
2. Vad är sannolikheten att vi på det åttonde kastet får vår andra sexa?
3. Vad är sannolikheten att vi fått exakt två sexor efter åtta kast?
4. Vad är väntevärdet och variansen av antalet kast som behövs för att få två sexor?

Hypergeometrisk fördelning

Antag att vi har N objekt, varav r har en egenskap vi är intresserade av. Vi väljer n objekt slumpmässigt utan återläggning och låter X vara antalet av de valda objekten som har egenskapen.



Då har X en hypergeometrisk fördelning med parametrar N , n och p , och motsvarande sannolikhetsfunktion ges av

$$f(k) = P(X = k) = \frac{\binom{r}{k} \binom{N-r}{n-k}}{\binom{N}{n}}$$

Obs! Om vi lägger tillbaka objekten blir $X \sim \text{bin}(n, r/N)$.

Hypergeometrisk fördelning

$$f(k) = P(X = k) = \frac{\binom{r}{k} \binom{N-r}{n-k}}{\binom{N}{n}}$$

Väntevärde och varians

$$\mathbb{E}[X] = \frac{nr}{N}$$

$$\text{Var}(X) = \frac{nr(N-r)(N-n)}{N^2(N-1)}$$

Hypergeometrisk fördelning

Exempel

Antag att man har testat en grupp med 20 personer för SARS-Cov-2, och att 5 av dem hade ett positivt testresultat. Antag att 6 slumpvis valda personer ur gruppen träffas på ett kafe, och låt X vara antalet bland dem har viruset.

1. Bestäm sannolikhetsfunktionen till X .
2. Beräkna väntevärde och varians till X .
3. Beräkna sannolikheten att man en av de som träffas har viruset.

Poissonfördelning

Poissonfördelningen räknar antalet händelser X som inträffar under ett tidsintervall om de inträffar oberoende av varandra, om det förväntade antalet händelser $\lambda > 0$ är känt.

Om n är stort och $np = \lambda$, så borde vi ha $X \approx \text{bin}(n, p)$, dvs $X \approx \text{bin}(n, \lambda/n)$, så

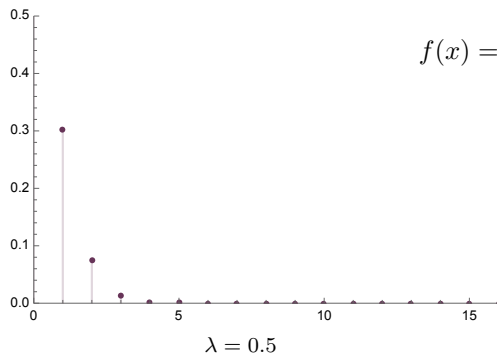
$$\begin{aligned} P(X = k) &\sim \binom{n}{k} (\lambda/n)^k (1 - \lambda/n)^{n-k} \sim \frac{n^k}{k!} (\lambda/n)^k (1 - \lambda/n)^n \\ &= \frac{\lambda^k}{k!} (1 - \lambda/n)^n \sim \frac{\lambda^k}{k!} e^{-\lambda} = \frac{e^{-\lambda} \lambda^k}{k!} \end{aligned}$$

Sannolikhetsfunktionen för en Poissonfördelning ges alltså av

$$f(x) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k \in \{1, 2, \dots\}$$

λ beskriver antalet händelser per tidsenhet, och kallas också för intensiteten för X . Vi skriver $X \sim \text{Poisson}(\lambda)$.

Poissonfördelning

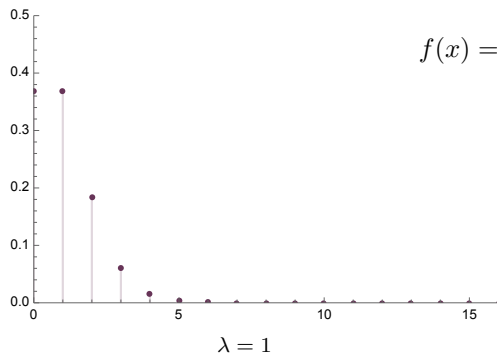


$$f(k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k \in \{0, 1, \dots\}$$

Väntevärde, varians och momentgenerande funktion

$$\mathbb{E}[X] = \lambda, \quad \text{Var}(X) = \lambda \quad \text{och} \quad m_X(t) = e^{\lambda(e^t - 1)}$$

Poissonfördelning

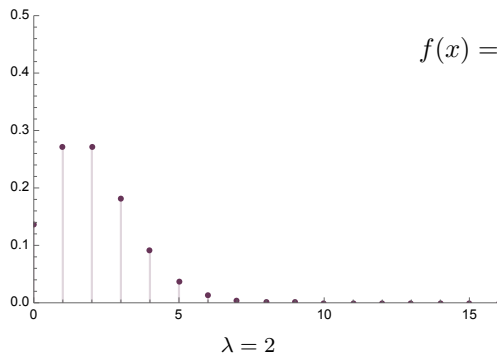


$$f(k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k \in \{0, 1, \dots\}$$

Väntevärde, varians och momentgenerande funktion

$$\mathbb{E}[X] = \lambda, \quad \text{Var}(X) = \lambda \quad \text{och} \quad m_X(t) = e^{\lambda(e^t - 1)}$$

Poissonfördelning

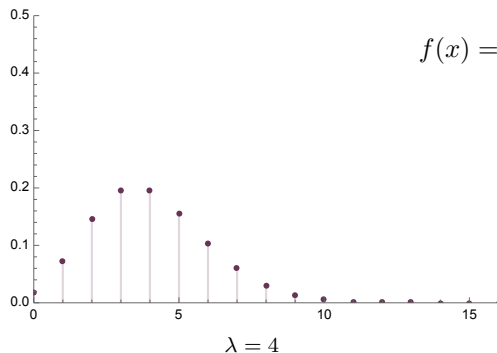


$$f(x) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k \in \{0, 1, \dots\}$$

Väntevärde, varians och momentgenerande function

$$\mathbb{E}[X] = \lambda, \quad \text{Var}(X) = \lambda \quad \text{och} \quad m_X(t) = e^{\lambda(e^t - 1)}$$

Poissonfördelning

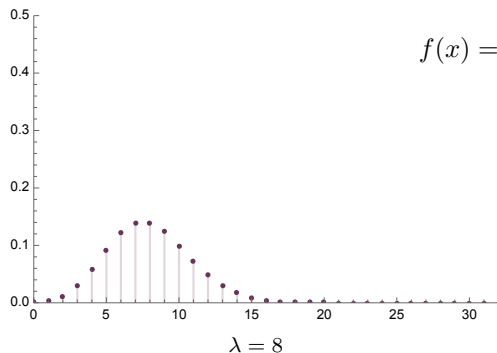


$$f(x) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k \in \{0, 1, \dots\}$$

Väntevärde, varians och momentgenerande function

$$\mathbb{E}[X] = \lambda, \quad \text{Var}(X) = \lambda \quad \text{och} \quad m_X(t) = e^{\lambda(e^t - 1)}$$

Poissonfördelning



$$f(x) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k \in \{0, 1, \dots\}$$

Väntevärde, varians och momentgenerande function

$$\mathbb{E}[X] = \lambda, \quad \text{Var}(X) = \lambda \quad \text{och} \quad m_X(t) = e^{\lambda(e^t - 1)}$$

Poissonfördelning

Poissonfördelningen räknar antalet händelser X som inträffar under ett tidsintervall om de inträffar oberoende av varandra, om det förväntade antalet händelser $\lambda > 0$ är känt.

Summa av oberoende Poissonfördelade stokastiska variabler

$$\begin{cases} X_1 \sim \text{Poisson}(\lambda_1) \\ X_2 \sim \text{Poisson}(\lambda_2) \end{cases} \Rightarrow X_1 + X_2 \sim \text{Poisson}(\lambda_1 + \lambda_2).$$

Poissonfördelning

Exempel

Låt X vara antalet tryckfel per sida om i en bok med i snitt 3 fel per sida. Antal att tryckfel händer oberoende av varandra.

1. Beräkna sannolikheten att en slumpvist vald sida innehåller minst ett tryckfel.
2. Var är sannolikheten att två slumpvist valda sidor tillsammans har fler än åtta tryckfel?

Exponentialfördelning

Anta att antalet händelser i intervallet $[0, 1]$ är $X \sim \text{Poisson}(\lambda)$. Låt T vara tiden vi får vänta tills någonting händer.

Låt $t > 0$ och låt X_t vara antalet händelser i intervallet $[0, t]$. Då är $X_t \sim \text{Poisson}(\lambda t)$.

Vi får

$$\begin{aligned} F_T(t) &= P(T \leq t) = P(X_t > 0) = 1 - P(X_t = 0) \\ &= 1 - \frac{e^{-\lambda t} \lambda^0}{0!} = 1 - e^{-\lambda t}. \end{aligned}$$

Vi säger att T är exponentialfördelad med parameter λ och skriver $T \sim \exp(\lambda)$.

Exponentialfördelning

Fördelningsfunktion och täthetsfunktion

$$F(t) = P(T \leq t) = 1 - e^{-\lambda t} \quad \text{och} \quad f(t) = F'(t) = \lambda e^{-\lambda t}, \quad t > 0.$$

Väntevärde, varians och momentgenererande funktion

$$\mathbb{E}[T] = \frac{1}{\lambda}, \quad \text{Var}(T) = \frac{1}{\lambda^2} \quad \text{och} \quad m_T(t) = \frac{1}{1 - t/\lambda}.$$

Exponentialfördelning

Minneslöshet

$$\begin{aligned}P(T \geq s + t \mid X \geq t) &= \frac{P(T \geq s + t \text{ och } T \geq t)}{P(T \geq t)} = \frac{P(T \geq s + t)}{P(T \geq t)} \\&= \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = e^{-\lambda s} = P(T \geq s).\end{aligned}$$

Koppling till Poissonfördelningen

Om $X \sim \text{Poisson}(\lambda)$, $X_t \sim \text{Poisson}(\lambda t)$, och $T_1, T_2, \dots \sim \exp(\lambda)$ så är

$$P(T_1 > t) = P(X_t = 0)$$

$$P(X \leq k) = P(T_1 + \dots + T_k \leq 1)$$

Mängden $\{T_1, T_1 + T_2, T_1 + T_2 + T_3, \dots\}$ kallas för en Poissonprocess med intensitet λ .

Exponentialfördelning

Exempel

Anta att antalet studenter som besöker kurshemsidan under en dag kan beskrivas med en Poissonfördelning, och att i snitt 30 studenter besöker kurshemsidan per timme. Vad är sannolikheten att det dröjer minst fem minuter innan någon student besöker kurshemsidan?

χ^2 -fördelning

Om $X_1, \dots, X_n \sim N(0, 1)$ är oberoende, så sägs $Y = X_1^2 + \dots + X_n^2$ ha en χ^2 -fördelning med n -frihetsgrader, och vi skriver $Y \sim \chi_n^2$.

Förväntat värde och varians

$$\mathbb{E}[Y] = \mathbb{E}[X_1^2 + \dots + X_n^2] = \mathbb{E}[X_1^2] + \dots + \mathbb{E}[X_n^2] = n \cdot 1 = n$$

$$\begin{aligned}\text{Var}(Y) &= \text{Var}(X_1^2 + \dots + X_n^2) = \text{Var}(X_1^2) + \dots + \text{Var}(X_n^2) \\ &= n \text{Var}(X_1^2) = n(\mathbb{E}[X^4] - \mathbb{E}[X^2]^2) = n(3 - 1) = 2n\end{aligned}$$

Fördelningsfunktion

Tabell IV på sidan 695-696 ger numeriska värden till fördelningsfunktionen $F(x)$ för χ^2 -fördelningen med olika antal frihetsgrader.

Gamma-fördelning

Gamma-funktionen

Gamma-funktionen är funktionen

$$\Gamma(\alpha) = \int_0^{\infty} z^{\alpha-1} e^{-z} dz$$

Man kan kontrollera att

- $\Gamma(1) = 1$
- $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$
- $\Gamma(n) = (n - 1)!$ för heltal n .

Gamma-fördelning

Slumpvariabeln X med täthetsfunktion

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

sägs vara Γ -fördelad med parametrar $\alpha, \beta > 0$. Vi skriver $X \sim \Gamma(\alpha, \beta)$.
Om $X \sim \Gamma(\alpha, \beta)$ så är

$$\mathbb{E}[X] = \alpha\beta, \quad \text{Var}(X) = \alpha\beta^2 \quad \text{och} \quad m_X(t) = (1 - \beta t)^{-\alpha}.$$

Specialfall

- Om $\alpha = 1$ och $\beta = 1/\lambda$ så är $\Gamma(1, 1/\lambda) = \exp(\lambda)$
- Om $\alpha = n/2$ och $\beta = 2$ så är $\Gamma(n/2, 2) = \chi_n^2$

Sammanfattning

