Föreläsning 4

Särskilda diskreta fördelningar

Exempel Låt oss påminna om fallet *dragning med återläggning*: Låt

$$X : \#\,\text{vita kulor i urvalet om } n \text{ kulor}$$

och

$$p = \frac{v}{v+s}\,,$$

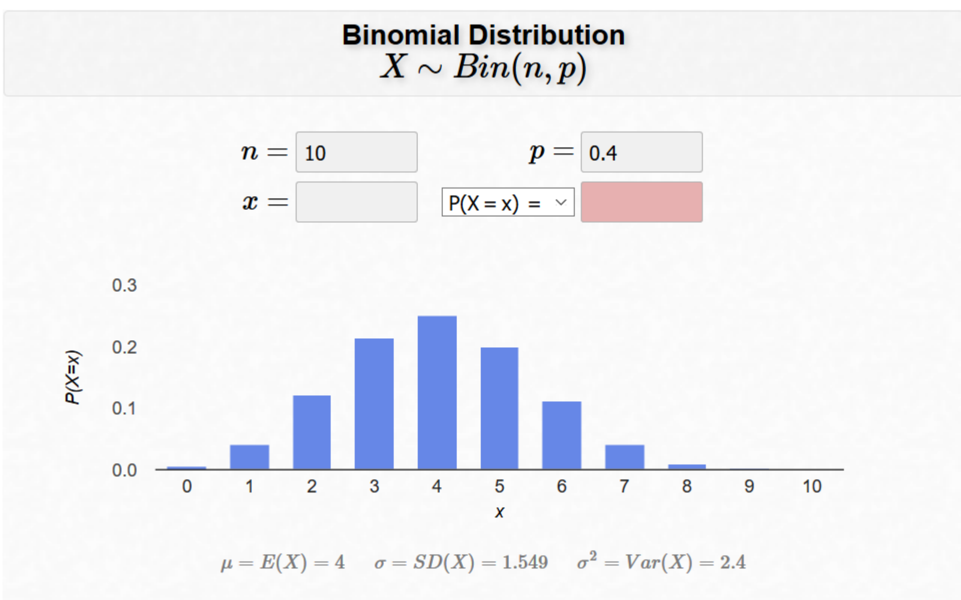
dvs. p är sannolikheten att dra en vit kula i en enda dragning. Vi visade att

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

Vi kan allmänt betrakta ett försök, där en händelse A kan inträffa med sannolikhet $p = P(A)$. Låt X vara antalet gånger som A inträffar i n oberoende upprepningar av detta försök.

Definition En diskret stokastisk variabel sägs vara *binomialfördelad* med parametrarna n och p ($Bin(n,p)$-fördelad) om

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



Matt Bognar U Iowa

Exempel Den amerikanska byggmarknaden *Homedepot* säljer skruvpaket med 10 skruvar. Från tidigare erfarenhet vet man att en skruv är felaktig med sannolikhet 0.01, oberoende av varandra. Homedepot erbjuder *pengarna tillbaka* garanti om två eller mer skruvar är felaktiga. Vilken andel av paketen måste Homedepot ersätta?

Google Bild

Svar: Låt

$$X : \#\text{ felaktiga skruvar i ett paket}\,.$$

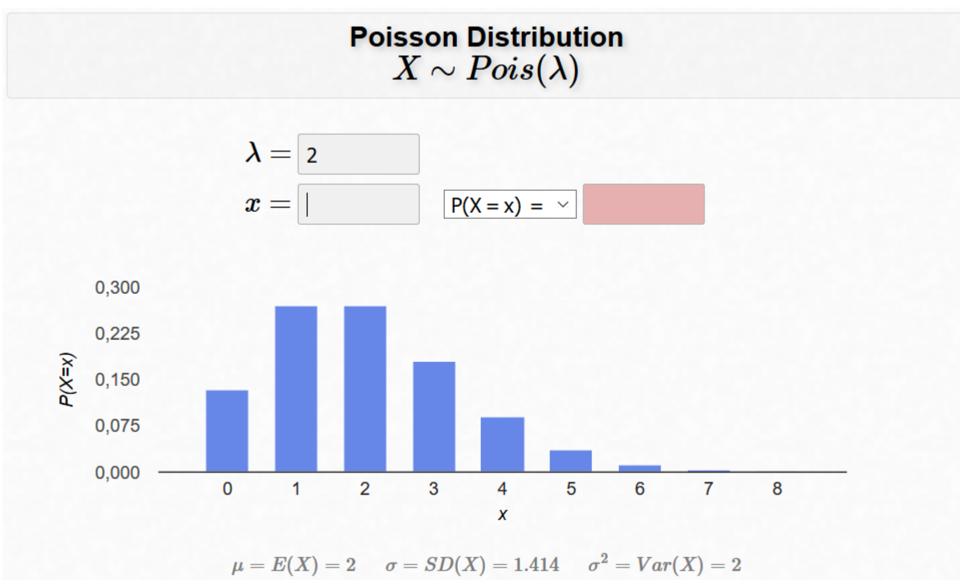
Då är $X \sim Bin(10, 0.01)$. Sökt är $P(X \geq 2)$.

$$\begin{aligned} P(X \geqslant 2) &= 1 - P(X=0) - P(X=1) \\ &= 1 - \binom{10}{0} 0.01^0 \cdot 0.99^{10} - \binom{10}{1} 0.01^1 \cdot 0.99^9 \\ &= 0.004\,, \end{aligned}$$

dvs. 0.4% au skruvpaketen måste Homdepot ersätta.

Definition En diskret stokastisk variabel sägs vara *Poisson fördelad* med parameter $\lambda > 0$ ($Po(\lambda)$-fördelad) om

$$p_X(k) = P(X = k) = \frac{\lambda^k}{k!} \cdot e^{-\lambda}\,, \quad k = 0, 1, \ldots\,.$$



Matt Bognar U Iowa

Egenskap Om n stort och p litet så kan man approximera

$$Bin(n, p) \approx Po(\lambda) \quad \text{där} \quad \lambda = np.$$

Vad betyder *n stort, p litet*?

Tumregel Vi använder oss av denna approximation om det inte finns någon tabell för $Bin(n,p)$ i vår tabellsamling, men det finns en tabell för $Po(\lambda)$ där $\lambda = np$.

Approximationen motiveras genom följande: Låt $X \sim Bin(n,p)$. Det gäller att

$$\begin{aligned}
p_X(k) = P(X=k) &= \binom{n}{k} p^k (1-p)^{n-k} \\
&= \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\
&= n \cdot (n-1) \cdot \ldots \cdot (n-k+1) \cdot \frac{1}{k!} \cdot \left(\frac{\lambda}{n}\right)^k \cdot \left(1 - \left(\frac{\lambda}{n}\right)\right)^{n-k} \\
&= \quad \frac{n}{n} \cdot \quad \frac{n-1}{n} \cdot \quad \ldots \cdot \quad \frac{n-k+1}{n} \cdot \quad \frac{\lambda^k}{k!} \cdot \quad \left[\left(1 - \left(\frac{\lambda}{n}\right)\right)^n\right]^{\frac{n-k}{n}}
\end{aligned}$$

$$\begin{array}{lcccccc}
 & \downarrow & \downarrow & & \downarrow & = & \downarrow \\
\text{då } n \longrightarrow \infty & 1 & 1 & \cdots & 1 & \dfrac{\lambda^k}{k!} & e^{-\lambda}
\end{array}$$

$$= \frac{\lambda^k}{k!} \cdot e^{-\lambda}, \quad k = 0, 1, 2, \ldots .$$

Exempel *Homedepot*-paket med 100 skruvar, *Pengarna tillbaka* garanti om 5 eller mer är felaktiga. Vilken andel måste Homedepot ersätta?

Google Bild

Svar: $X \sim Bin(100, 0.01)$ approximeras med $Po(1)$. Ur tabellen fås $P(X \geq 5) = 0.0037$.

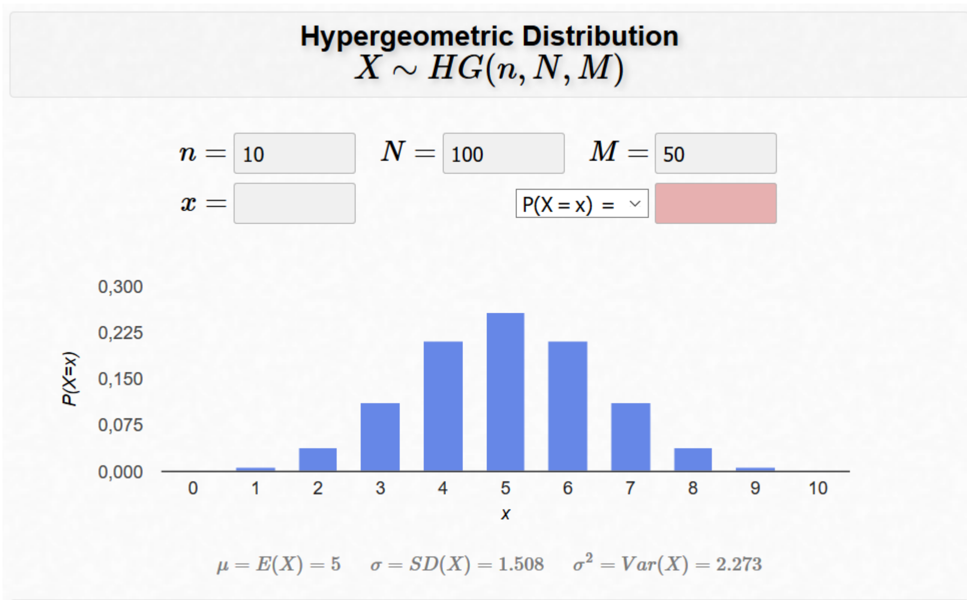
Definition En diskret stokastisk variabel X sägs vara *hypergeometrisk* fördelad med parametrarna N, n, p ($Hyp(N,n,p)$-fördelad) om med

$$N = v + s \quad \text{och} \quad p = \frac{v}{v+s}$$

det gäller att

$$p_X(k) = P(X=k) = \frac{\binom{v}{k} \cdot \binom{s}{n-k}}{\binom{v+s}{n}}, \quad k = 0, \ldots, v, \quad n-k = 0, \ldots, s.$$

I denna bild: $HG(n, N, M)$ motsvarar $Hyp(N, n, p)$ med $p = M/N$



Matt Bognar U Iowa

Exempel En urna innehåller 4 vita och 6 svarta kulor. Man drar 3 kulor. Låt

$$X : \# \text{ vita bland de dragna.}$$

Vad är sannolikheten att man drar fler vita än svarta kulor, utan återläggning?

Svar: Sökt är

$$\begin{aligned}
P(X \geq 2) &= P(X = 2) + P(X = 3) \\
&= \frac{\binom{4}{2} \cdot \binom{6}{1}}{\binom{10}{3}} + \frac{\binom{4}{3} \cdot \binom{6}{0}}{\binom{10}{3}} = \frac{3}{10} + \frac{1}{30} \\
&= \frac{1}{3}.
\end{aligned}$$

Väntevärde och varians

Definition Låt X vara en diskret stokastisk variabel.
(a)

$$E[X] = \sum_{\text{alla möjliga värden } k \text{ på } X} k\, P(X = k)$$

kallas *väntevärde* av X.
(b)

$$Var(X) \equiv V(X) = E\left[(X - E[X])^2\right]$$

$$= \sum_{\text{alla möjliga värden } k \text{ på } X} (k - E[X])^2 \, P(X = k)$$

kallas *varians* av X.
(c)

$$D(X) = \sqrt{Var(X)}$$

kallas *standardavvikelse* av X.

Anmärkningar (1) Vi säger att väntevärdet *existerar* om högra sidan är absolut summerbar, dvs.

$$\sum_{\text{alla } k} |k| \, P(X = k) < \infty.$$

(2) Ofta fördelaktigt att använda den så kallade *förkortningsformeln*,

$$Var(X) = E[X^2] - (E[X])^2$$

$$= \sum_{\text{alla } k} k^2 \cdot P(X = k) - (E[X])^2 \, .$$

Sats (Räkneregler) (a) $E[a] = a$, $Var(a) = 0$ för varje icke-slumpmässigt $a \in \mathbb{R}$.
(b)

$$E[aX + b] = aE[X] + b \, , \quad a, b \in \mathbb{R},$$

$$Var(aX + b) = a^2 Var(X) \, , \quad a, b \in \mathbb{R}, \quad \text{(inget } b \text{ å högra sidan).}$$

(c) Låt g vara en reell funktion. Det gäller att

$$E[g(X)] = \sum_{\text{alla } k} g(k) \, P(X = k)$$

om högra sidan är absolut summerbar, dvs. $\sum_{\text{alla } k} |g(k)| \, P(X = k) < \infty$.

Anmärkningar (1) Med (c) bevisas andra delen av förkortningsformeln. Här tas $g(x) = x^2$, $x \in \mathbb{R}$.
(2) Vi säger att variansen och standardavvikelsen *existerar* om

$$\sum_{\text{alla } k} k^2 \, P(X = k) < \infty.$$

Detta medför att väntevärdet existerar.

Exempel Tärningskast

$$E[X] = 1 \cdot P(X = 1) + 2 \cdot P(X = 2) + \ldots + 6 \cdot P(X = 6)$$
$$= 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \cdots + 6 \cdot \frac{1}{6} = 3.5 \, .$$

Dessutom

$$\begin{aligned} Var(X) &= E[X^2] - (E[X])^2 \quad \text{(se förkortningsformel)} \\ &= 1^2 \cdot P(X=1) + 2^2 \cdot P(X=2) + \ldots + 6^2 \cdot P(X=6) - 3.5^2 \\ &= \frac{35}{12}\,. \end{aligned}$$

Sats (a) Låt $X \sim Bin(n,p)$. Det gäller att

$$E[X] = np \quad \text{och} \quad Var(X) = np(1-p).$$

(b) Låt $X \sim Po(X)$. Det gäller att

$$E[X] = \lambda \quad \text{och} \quad Var(X) = \lambda.$$

(c) Låt $X \sim Hyp(N,n,p)$. Det gäller att

$$E[X] = np \quad \text{och} \quad Var(X) = np(1-p)\,\frac{N-n}{N-1}.$$

Exempel *Homedepot* säljer skruvpaket med 100 skruvar. Var och en skruv är felaktig med sannolikhet 0.01.

(a) Man väljer ut 10 skruvar för kontroll. Man granskar en skruv och stoppar den tillbaka till urvalet efter att man har noterat om skruven är fungerande eller felaktig. Vad är det förväntade antalet felaktiga skruvar i urvalet?
(b) Man väljer ut 10 skruvar på en gång. Vad är det förväntade antalet felaktiga skruvar i urvalet?

Google Bild

Svar: Låt

$$X : \#\ \text{felaktiga skruvar i urvalet.}$$

(a) Då är $X \sim Bin(10, 0.01)$, dvs. $n = 10$ och $p = 0.01$. Således

$$E[X] = np = 0.1\,.$$

(b) Här har vi $X \sim Hyp(100, 10, 0.01)$, dvs. $N = 100$, $n = 10$, $p = 0.01$. Således

$$E[X] = np = 0.1\,.$$

Exempel Partiklar emitteras från en radioaktiv kropp med intensitet 10 partiklar per sekund. Vi får anta att antalet emitterade partiklar per sekund X följer en Poissonfördelning.

(a) Bestäm parametern λ.
(b) Vad är sannolikheten att högst 10 partiklar emitteras under en viss sekund?
(c) Vad är sannolikheten att minst 10 partiklar emitteras under en viss sekund?

Svar: (a) $\lambda = E[X] = 10$ partiklar per sekund. Alltså $X \sim Po(10)$.
(b) Sökt är $P(X \leq 10) = 0.583$ från tabellen.
(c) Sökt är $P(X \geq 10) = 0.542$ från tabellen.

Jörg-Uwe Löbus, jorg-uwe.lobus@liu.se