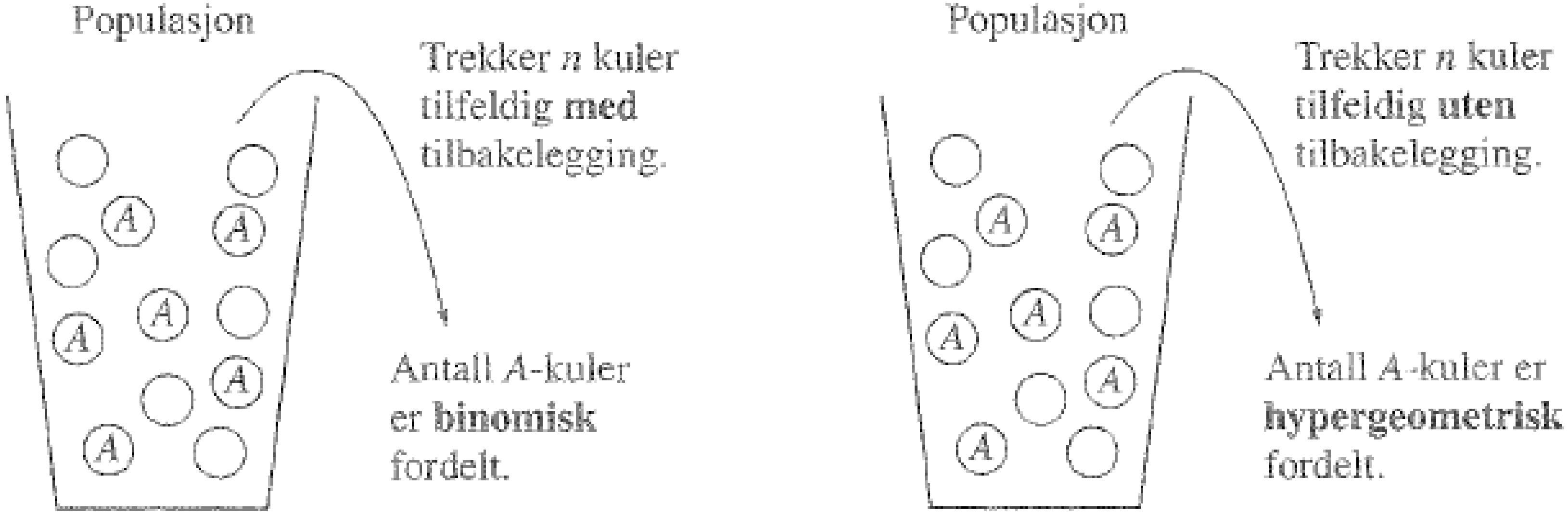
ST0103 Brukerkurs i statistikk
Høst 2016

Forelesning 12, 30.9.2016

5.3 Hypergeometrisk fordeling



Figur 5.5 Urnemodellen forklarer forskjellen på binomisk og hypergeometrisk fordeling

5.4 Geometrisk fordeling

Regel 5.6 (Geometrisk fordeling) Y er geometrisk fordelt med parameter p hvis

$$P(Y = y) = p \cdot (1 - p)^{y-1}, \qquad \text{for} \quad y = 1, 2, \ldots$$

En geometrisk variabel har forventning og varians

$$E(Y) = \frac{1}{p}, \qquad \mathrm{Var}(Y) = \frac{1 - p}{p^2}$$

LOTTO

Trekning: 24.09.2016

Vinnertall

1 2 7 10 16 21 27

Tilleggstall

11

Premieoversikt

Antall rette	Premie	Vinnere
7	3 586 940 kr	4
6 + 1	155 760 kr	13
6	3 425 kr	619
5	95 kr	23 144
4	45 kr	315 583

Omsetning 77 138 530 kr

Teori vs. empiri

X = antall rekker med 7 rette i Lotto.

Med totalt $n = 1542706$ rekker (24.09.2016) blir $E(X) = \mu = 2.868$

Antar dermed X er Poisson-fordelt med $\mu = 2.868$.

Sammenligner med empirisk relativ frekvens observert i de første 39 uker av 2016 (hvor antall rekker har vært tilnærmet konstant).

x	$P(X = x) = \frac{\mu^x}{x!}e^{-\mu}$	Relativ frekvens 2016	Frekvens 2016 (forventet)
0	0.057	0.026	1 (2.2)
1	0.163	0.179	7 (6.4)
2	0.234	0.256	10 (9.1)
3	0.223	0.179	7 (8.7)
4	0.160	0.205	8 (6.2)
5	0.092	0.051	2 (3.6)
6	0.044	0.077	3 (1.7)
7	0.018	0.026	1 (0.7)
8+	0.009	0	0 (0.3)

Eksempel – Poissontilnærmelsen

La oss se hvordan tilnærmelsen virker på et konkret eksempel. Vi lar X være binomisk fordelt med parametre $n = 100$ og $p = 0.02$. Denne fordelingen kan tilnærmes med Poissonfordelingen med parameter $\lambda = n\,p = 100 \cdot 0.02 = 2.0$.

$$\mathrm{P}(X = x) \;=\; \binom{100}{x} 0.02^x \; 0.98^{100-x} \;\approx\; \frac{2.0^x}{x!}\, e^{-2.0} \qquad ;\; x \;=\; 0, 1, 2, 3, \ldots$$

De eksakte og de tilnærmede punktsannsynlighetene blir som følger (Minitab):

x	0	1	2	3	4	5	6	7	8
$\mathrm{P}(X = x)$ (bin)	0.133	0.271	0.273	0.182	0.090	0.035	0.011	0.003	0.001
$\mathrm{P}(X = x)$ (Po)	0.135	0.271	0.271	0.180	0.090	0.036	0.012	0.003	0.001

For $x \geq 9$ har vi $\mathrm{P}(X = x) = 0.000$ både i eksakt og i tilnærmet fordeling. Som vi ser er Poissontilnærmelsen meget god for disse verdiene av n og p. ■

APRIL 1999 33:4 ANNALS OF EMERGENCY MEDICINE

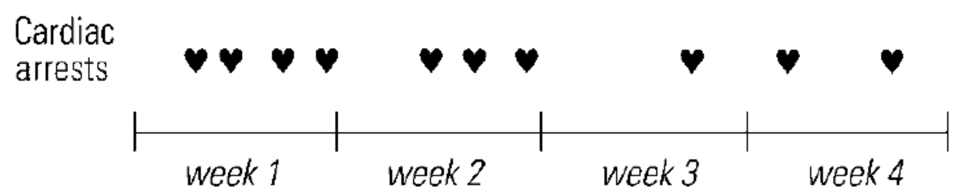
Modeling the Occurrence of Cardiac Arrest as a Poisson Process

Eirik Skogvoll, MD*

Bo Henry Lindqvist, PhD‡

Figure 1.

Modeling cardiac arrests along the time axis. Ten hypothetical events of cardiac arrest (♥) are plotted along a 4-week time axis, illustrating the Poisson random variable X (count of cardiac arrests per week) and the exponentially distributed variable T (interval between consecutive events). The Poisson parameter estimate in this example is 2.5 events per week ($\overline{x}$, the mean event rate).

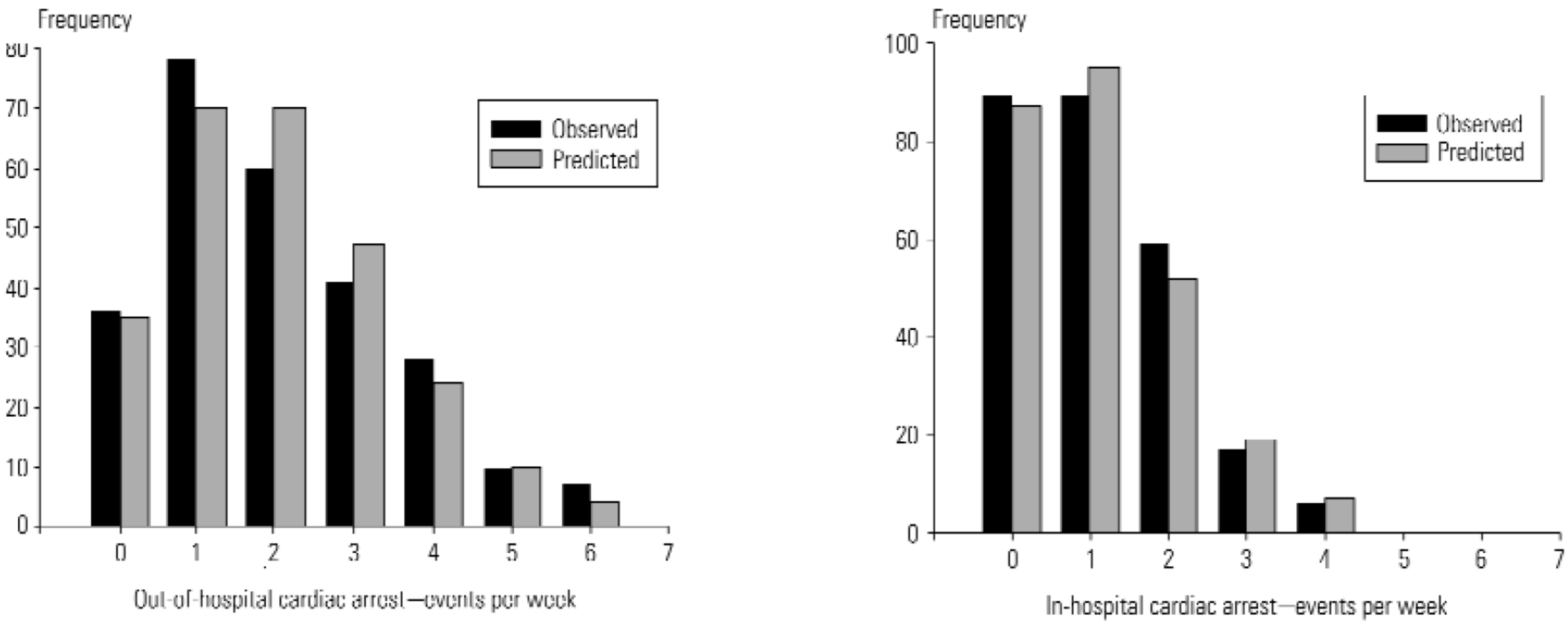


Observations

Interval (t, hours)	26	47	37	79	42	47	147	84	95	
♥per week (X, count)	4			3			1		2	($\overline{x}$ = 2.5)

Figure 2.

Parameter estimates and observed versus predicted frequency distributions of the weekly counts of cardiac arrest during the 5-year investigation period.



Mean weekly event rate:	2.019
Sample variance:	2.227
Calculated mean interval:	83 h
Observed SD (interval):	86 h

Mean weekly event rate:	1.092
Sample variance:	1.098
Calculated mean interval:	154 h
Observed SD (interval):	153 h

Predicted = ved bruk av Poisson-fordeling

Mean interval = forventet tid mellom hendelser (eksponentialfordelt)