NTNU
Kunnskap for ei betre verd

Negativ binomisk og Poissonfordelinga

Torstein Fjeldstad
Institutt for matematiske fag, NTNU
07.02.2019

I dag

- Geometrisk og negativ binomisk fordeling
- Poissonfordeling

Repetisjon

Repetisjon I

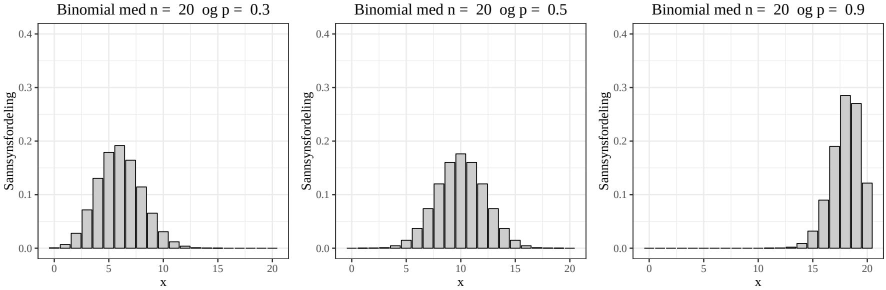
Krav Bernoulli-prosess

1. n uavhengige forsøk
2. I kvart forsøk i har me suksess, $I_i = 1$, eller ikkje-suksess, $I_i = 0$
3. Sannsynet for suksess er konstant for alle n forsøka;
   $p = P(I_i = 1)$ for alle $i = 1, \ldots, n$

Binomisk fordeling

La $X = \sum_{i=1}^{n} I_i$ vere talet på suksessar i ein Bernoulli-prosess. Då er X binomisk fordelt med

$$P(X = x) = b(x; n, p) = \binom{n}{x} p^x (1 - p)^{n-x}.$$


Binomial med n = 20 og p = 0.3
Binomial med n = 20 og p = 0.5
Binomial med n = 20 og p = 0.9
Sannsynsfordeling
x
0.0
0.1
0.2
0.3
0.4
0
5
10
15
20

Repetisjon II

Krav multinomisk-prosess

1. n uavhengige forsøk
2. Kvart forsøk resulterar i nøyaktig ein av R kategoriar
3. Sannsynet for kvar kategori r er lik i kvart forsøk i:
   $P(\text{kategori } r) = p_r$

Multinomisk fordeling

Tala på utfall $X = (X_1, \ldots, X_R)$ i ein mulitnomisk-prosess er multinomisk fordelt med

$$f(x_1, \ldots, x_R; n, p_1, \ldots, p_R) = \binom{n}{x_1, \ldots, x_R} p_1^{x_1} \ldots p_R^{x_R}$$

der $\sum_{r=1}^R x_r = n$ og $\sum_{r=1}^R p_r = 1$.

Repetisjon III

Situasjon hypergeometrisk fordeling

- Urne med N kuler
- k blåe kuler (suksess)
- $N - k$ raude kuler (fiasko)
- Trekk n kuler (utan tilbakelegging)
- X er talet på dei trekte kulene som er blåe (talet på suksessar)

Hypergeometrisk fordeling

En hypergeometrisk stokastisk variabel X har fordeling

$$f(x; N, n, k) = h(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}.$$

Geometrisk og negativ binomisk fordeling

Geometrisk fordeling

Geometrisk fordeling

Dersom me har uavhengig forsøk, kvar med suksessannsyn p, og lar X vere talet på forsøk til og med første suksess er X geometrisk fordelt med:

$$f(x;p) = g(x;p) = (1-p)^{x-1}p \qquad x = 1, 2, 3, \ldots$$

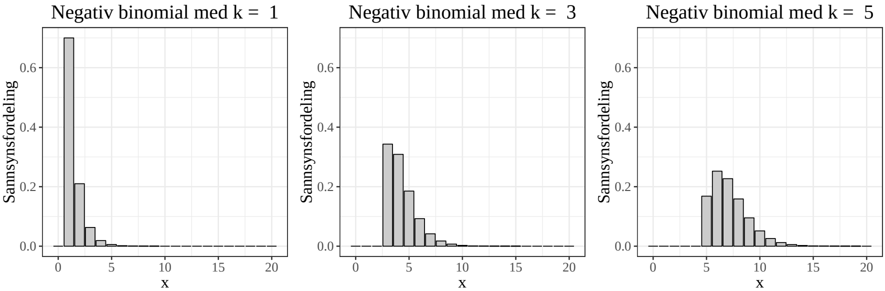
.

Negativ binomisk fordeling

Negativ binomisk fordeling

Dersom me har uavhengige forsøk, kvar med suksessannsyn p, og lar X vere talet på forsøk til og med k-te sukesess, så er X negativ binomisk fordelt med

$$f(x;k,p) = b^\star(x;k,p) = \binom{x-1}{k-1} p^k (1-p)^{x-k} \qquad x = k, k+1, k+2, \ldots$$


Negativ binomial med k = 1
Negativ binomial med k = 3
Negativ binomial med k = 5
Sannsynsfordeling
0.0
0.2
0.4
0.6
0
5
10
15
20
x

Poissonfordelinga

Eksempel (buss)

Oppgåve

La X vere talet på bussar som passerer Samfundet i løpet av T minutt.

Kva er fordelinga til X?

Poissonfordelinga I

- Talet på hendingar i eit tidsintervall er uavhengig av talet på hendingar i disjunkte tidsintervall
- Sannsynet for at ei hending inntreff i eit kort tidsintervall er proporsjonalt med lengda på tidsintervallet

$$P(\text{"}X = 1\text{" i intervallet } [t, t + \Delta t)) = \lambda \Delta t + o(\Delta t) \rightarrow \lambda \Delta t$$

- Sannsynet for at det inntreff meir enn ei hending innanfor eit lite tidsintervall er neglisjerbart

$$P(\text{"}X \geq 2\text{" i intervallet } [t, t + \Delta t)) = o(\Delta t) \rightarrow 0$$

Merk: me kan byte ut tid med for eksempel distanse, areal, volum

Poissonfordelinga II

Poissonfordelinga

La X vere talet på hendingar i eit tidsintervall av lengde t. X er poissonfordelt med

$$f(x; \lambda t) = p(x; \lambda t) = \frac{(\lambda t)^x}{x!} \exp(-\lambda t)$$

der λ er gjennomsnittleg tal på hendingar per tidseining.

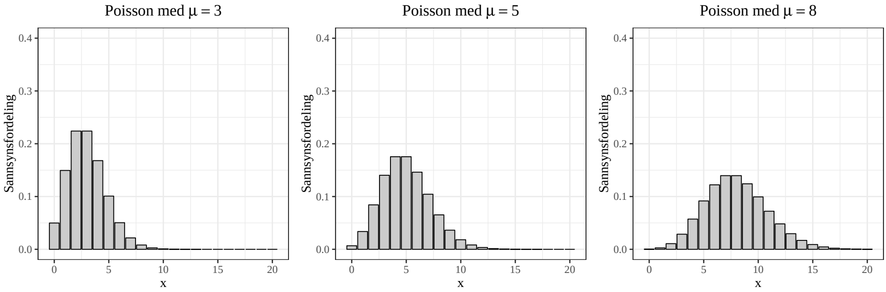
Forventningsverdi og varians

For ein poissonfordelt stokastisk variabel X med parameter λt er:

$$E(X) = \lambda t \qquad \text{og} \qquad Var(X) = \lambda t.$$

Merk me nyttar ofte $\mu = \lambda t$.

Poissonfordelinga III



Eksempel (buss) I

X = talet på busser som passerer Samfundet i løpet av 5 minutt
$\lambda = 3$ busser/min
$X \sim$ Poisson$(\lambda = 5 \cdot 3 = 15)$

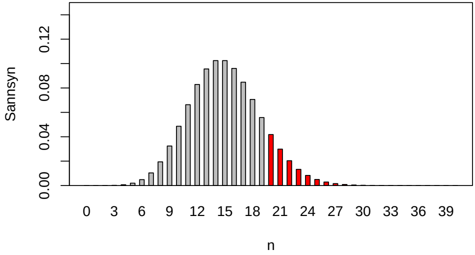


Figure: $P(X > 20)$

Eksempel (buss) II

Poissonfordeling

$P(X \leq x)$

$x \backslash \mu$	10	11	12	13	14	15	16	17	18
0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
1	.0005	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000
2	.0028	.0012	.0005	.0002	.0001	.0000	.0000	.0000	.0000
3	.0103	.0049	.0023	.0011	.0005	.0002	.0001	.0000	.0000
4	.0293	.0151	.0076	.0037	.0018	.0009	.0004	.0002	.0001
5	.0671	.0375	.0203	.0107	.0055	.0028	.0014	.0007	.0003
6	.1301	.0786	.0458	.0259	.0142	.0076	.0040	.0021	.0010
7	.2202	.1432	.0895	.0540	.0316	.0180	.0100	.0054	.0029
8	.3328	.2320	.1550	.0998	.0621	.0374	.0220	.0126	.0071
9	.4579	.3405	.2424	.1658	.1094	.0699	.0433	.0261	.0154
10	.5830	.4599	.3472	.2517	.1757	.1185	.0774	.0491	.0304
11	.6968	.5793	.4616	.3532	.2600	.1848	.1270	.0847	.0549
12	.7916	.6887	.5760	.4631	.3585	.2676	.1931	.1350	.0917
13	.8645	.7813	.6815	.5730	.4644	.3632	.2745	.2009	.1426
14	.9165	.8540	.7720	.6751	.5704	.4657	.3675	.2808	.2081
15	.9513	.9074	.8444	.7636	.6694	.5681	.4667	.3715	.2867
16	.9730	.9441	.8987	.8355	.7559	.6641	.5660	.4677	.3751
17	.9857	.9678	.9370	.8905	.8272	.7489	.6593	.5640	.4686
18	.9928	.9823	.9626	.9302	.8826	.8195	.7423	.6550	.5622
19	.9965	.9907	.9787	.9573	.9235	.8752	.8122	.7363	.6509
20	.9984	.9953	.9884	.9750	.9521	.9170	.8682	.8055	.7307
21	.9993	.9977	.9939	.9859	.9712	.9469	.9108	.8615	.7991
22	.9997	.9990	.9970	.9924	.9833	.9673	.9418	.9047	.8551
23	.9999	.9995	.9985	.9960	.9907	.9805	.9633	.9367	.8989
24	1.0000	.9998	.9993	.9980	.9950	.9888	.9777	.9594	.9317
25	1.0000	.9999	.9997	.9990	.9974	.9938	.9869	.9748	.9554
26	1.0000	1.0000	.9999	.9995	.9987	.9967	.9925	.9848	.9718
27	1.0000	1.0000	.9999	.9998	.9994	.9983	.9959	.9912	.9827
28	1.0000	1.0000	1.0000	.9999	.9997	.9991	.9978	.9950	.9897
29	1.0000	1.0000	1.0000	1.0000	.9999	.9996	.9989	.9973	.9941
30	1.0000	1.0000	1.0000	1.0000	.9999	.9998	.9994	.9986	.9967
31	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9997	.9993	.9982
32	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9996	.9990

Samanheng binomisk og poissonfordeling

Teorem

La $X \sim \text{binomisk}(x; n, p)$. Når $n \to \infty$ og $p \to 0$ slik at $\mu = np$ er konstant har me at

$$\text{binomisk}(x; n, p) \to \text{poisson}(x; \mu)$$

Neste veke

— Kontinuerlege fordelingar