ST0103 Brukerkurs i statistikk
Høst 2016

Forelesning 11, 28.9.2016

Definisjon 4.16 (Uavhengige variabler) To diskrete stokastiske variabler X og Y er uavhengige hvis og bare hvis følgende ligning er tilfredsstilt for *alle* mulige verdipar (x, y) i simultanfordelingen til X og Y.

$$P(x, y) = P(X = x) \cdot P(Y = y) \qquad [4.17]$$

Regel 4.17 (Varians til sum av uavhengige variabler) Hvis $X_1, X_2, \ldots, X_n$ er *uavhengige* stokastiske variabler, gjelder (a-ene og b er konstanter):

$$\begin{aligned}\operatorname{Var}(a_1X_1 + a_2X_2 + \cdots + a_nX_n + b)\\ = a_1^2 \operatorname{Var}(X_1) + a_2^2 \operatorname{Var}(X_2) + \cdots + a_n^2 \operatorname{Var}(X_n) \qquad [4.18]\end{aligned}$$

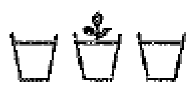
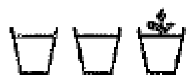
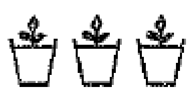
Regel 4.18 (Produkt av uavhengige variabler) Dersom $X_1, X_2, \ldots, X_n$ er *uavhengige* stokastiske variabler, gjelder:

$$E(X_1 \cdot X_2 \cdot \ldots \cdot X_n) = E(X_1) \cdot E(X_2) \cdot \ldots \cdot E(X_n) \qquad [4.19]$$

5.2 Binomisk modell

X = ant. frø som spirer

p = P(frø spirer)

Resultat	Kombinasjon	Sannsynlighet	Sum sannsynlighet ($p = 0.67$
Ingen frø spirer		$(1-p)\cdot(1-p)\cdot(1-p)$	$P(X=0)=(1-p)^3 \quad = 0.036$
Ett frø spirer		$p\cdot(1-p)\cdot(1-p)$	
		$(1-p)\cdot p\cdot(1-p)$	$P(X=1)=3p(1-p)^2=0.219$
		$(1-p)\cdot(1-p)\cdot p$	
To frø spirer		$p\cdot p\cdot(1-p)$	
		$p\cdot(1-p)\cdot p$	$P(X=2)=3p^2(1-p)=0.444$
		$(1-p)\cdot p\cdot p$	
Tre frø spirer		$p\cdot p\cdot p$	$P(X=3)=p^3 \quad = 0.301$

Figur 5.2 Et binomisk forsøk med planting av tre solsikkefrø

Binomisk forsøksrekke

Binomisk eksperiment: Et eksperiment som består i gjentatte forsøk med følgende egenskaper:

1. Det er n identiske uavhengige forsøk.
2. Hvert forsøk har to mulige utfall, ofte kalt suksess og fiasko.
3. P(suksess)=p, P(fiasko)=q, $p + q = 1$.
4. Den binomiske tilfeldige variabelen x er antallet suksessfulle utfall som inntreffer, og x kan anta enhver heltallsverdi fra 0 til n.

Forstå $P(x)$

$$P(x) = \binom{n}{x}(p^x)(q^{n-x}) \text{ for } x = 0, 1, 2, \ldots, n$$

Vi ser at sannsynlighetsfunksjonen er produktet av tre faktorer:

— På hvor mange måter kan akkurat x suksesser inntreffe i løpet av n forsøk: $\binom{n}{x}$.

— Sannsynligheten for akkurat x suksesser: p^x.

— Sannsynligheten for akkurat $n - x$ fiaskoer: $q^{(n-x)}$, der $q = 1 - p$.

Forventning og varians

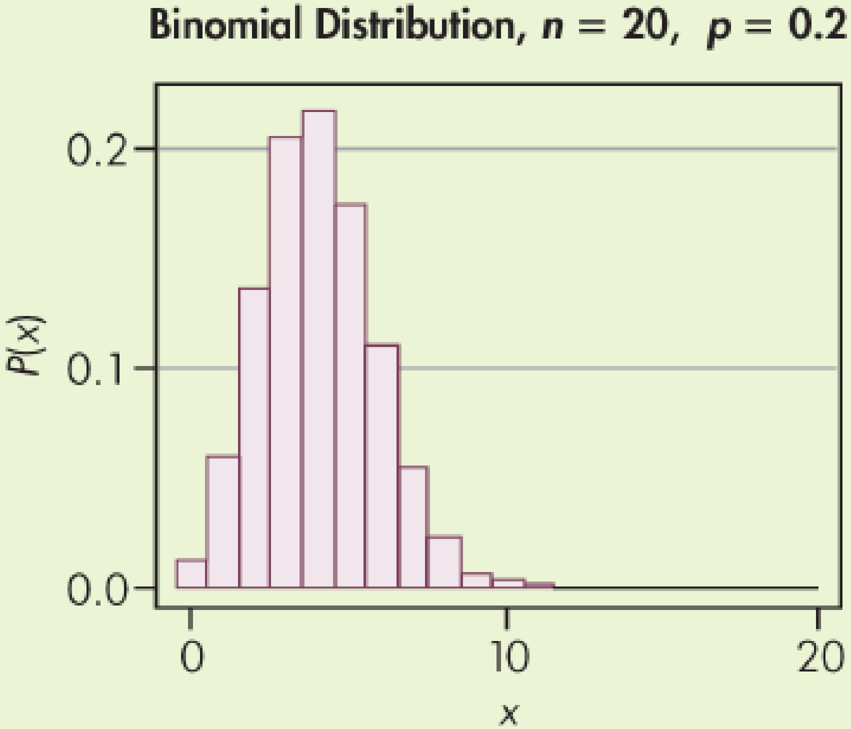
Regel 5.3 (Forventning og varians) En binomisk variabel X har forventningsverdi og varians lik

$$E(X) = np$$

$$\mathrm{Var}(X) = np(1-p)$$

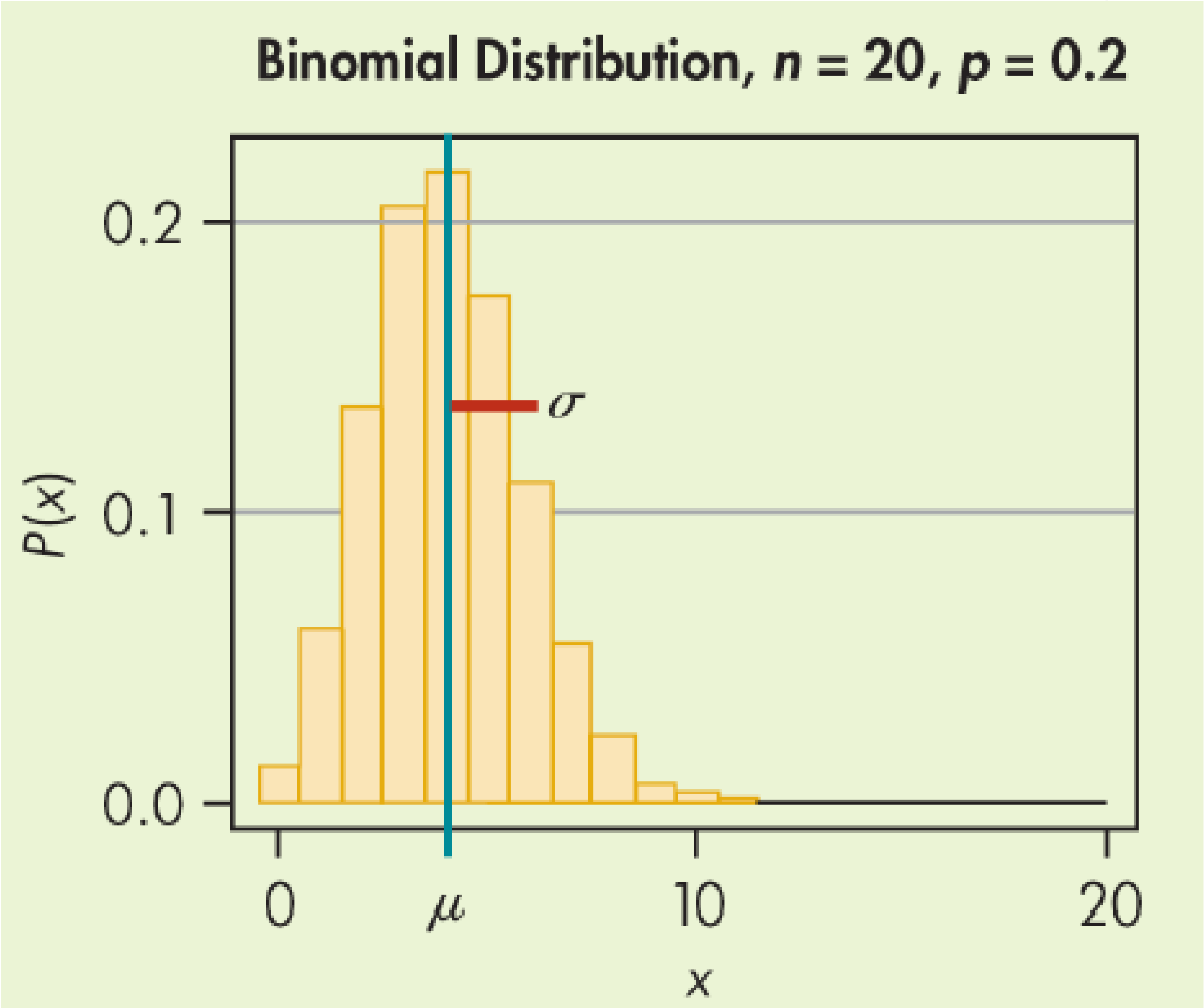
Grafisk fordeling

x	$P(x)$
0	0.012
1	0.058
2	0.137
3	0.205
4	0.218
5	0.175
6	0.109
7	0.055
8	0.022
9	0.007
10	0.002
11	0+
12	0+
13	0+
⋮	⋮
20	0+



Binomisk fordeling med $n = 20$ og $p = 0.2$.

Grafisk fordeling med μ og σ



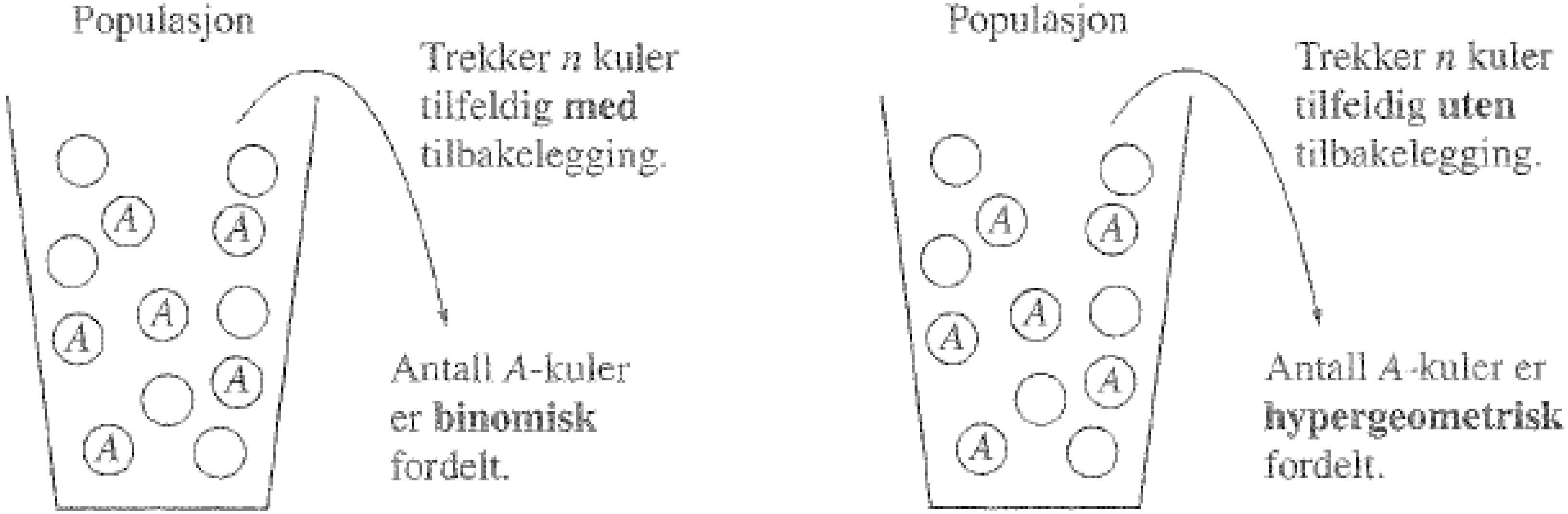
Binomisk fordeling med $n = 20$ og $p = 0.2$.

Binomisk fordeling

$$P(X \leq x)$$

		p											
n	x	.05	.10	.15	.20	.25	.30	.40	.50	.60	.70	.80	.90
9	0	.630	.387	.232	.134	.075	.040	.010	.002	.000	.000	.000	.000
	1	.929	.775	.599	.436	.300	.196	.071	.020	.004	.000	.000	.000
	2	.992	.947	.859	.738	.601	.463	.232	.090	.025	.004	.000	.000
	3	.999	.992	.966	.914	.834	.730	.483	.254	.099	.025	.003	.000
	4	1.000	.999	.994	.980	.951	.901	.733	.500	.267	.099	.020	.001
	5	1.000	1.000	.999	.997	.990	.975	.901	.746	.517	.270	.086	.008
	6	1.000	1.000	1.000	1.000	.999	.996	.975	.910	.768	.537	.262	.053
	7	1.000	1.000	1.000	1.000	1.000	1.000	.996	.980	.929	.804	.564	.225
	8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.998	.990	.960	.866	.613
10	0	.599	.349	.197	.107	.056	.028	.006	.001	.000	.000	.000	.000
	1	.914	.736	.544	.376	.244	.149	.046	.011	.002	.000	.000	.000
	2	.988	.930	.820	.678	.526	.383	.167	.055	.012	.002	.000	.000
	3	.999	.987	.950	.879	.776	.650	.382	.172	.055	.011	.001	.000
	4	1.000	.998	.990	.967	.922	.850	.633	.377	.166	.047	.006	.000
	5	1.000	1.000	.999	.994	.980	.953	.834	.623	.367	.150	.033	.002
	6	1.000	1.000	1.000	.999	.996	.989	.945	.828	.618	.350	.121	.013
	7	1.000	1.000	1.000	1.000	1.000	.998	.988	.945	.833	.617	.322	.070
	8	1.000	1.000	1.000	1.000	1.000	1.000	.998	.989	.954	.851	.624	.264
	9	1.000	1.000	1.000	1.000	1.000	1.000	1.000	.999	.994	.972	.893	.651
11	0	.569	.314	.167	.086	.042	.020	.004	.000	.000	.000	.000	.000
	1	.898	.697	.492	.322	.197	.113	.030	.006	.001	.000	.000	.000
	2	.985	.910	.779	.617	.455	.313	.119	.033	.006	.001	.000	.000
	3	.998	.981	.931	.839	.713	.570	.296	.113	.029	.004	.000	.000
	4	1.000	.997	.984	.950	.885	.790	.533	.274	.099	.022	.002	.000
	5	1.000	1.000	.997	.988	.966	.922	.753	.500	.247	.078	.012	.000
	6	1.000	1.000	1.000	.998	.992	.978	.901	.726	.467	.210	.050	.003
	7	1.000	1.000	1.000	1.000	.999	.996	.971	.887	.704	.430	.161	.019
	8	1.000	1.000	1.000	1.000	1.000	.999	.994	.967	.881	.687	.383	.090

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

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