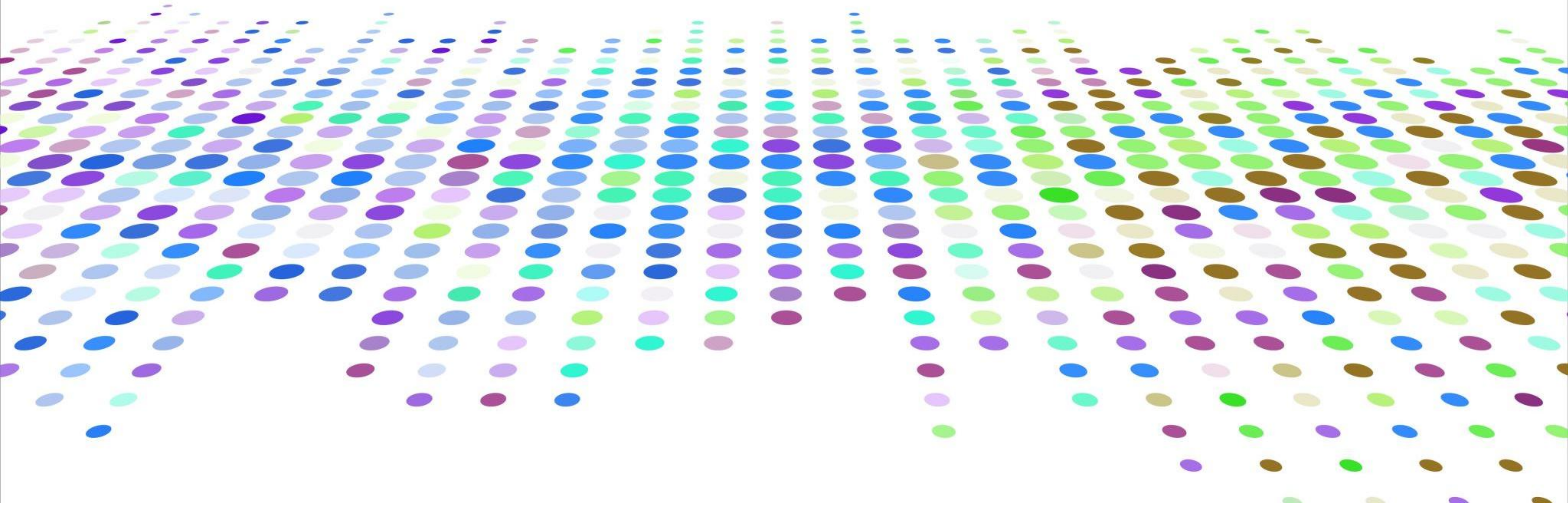
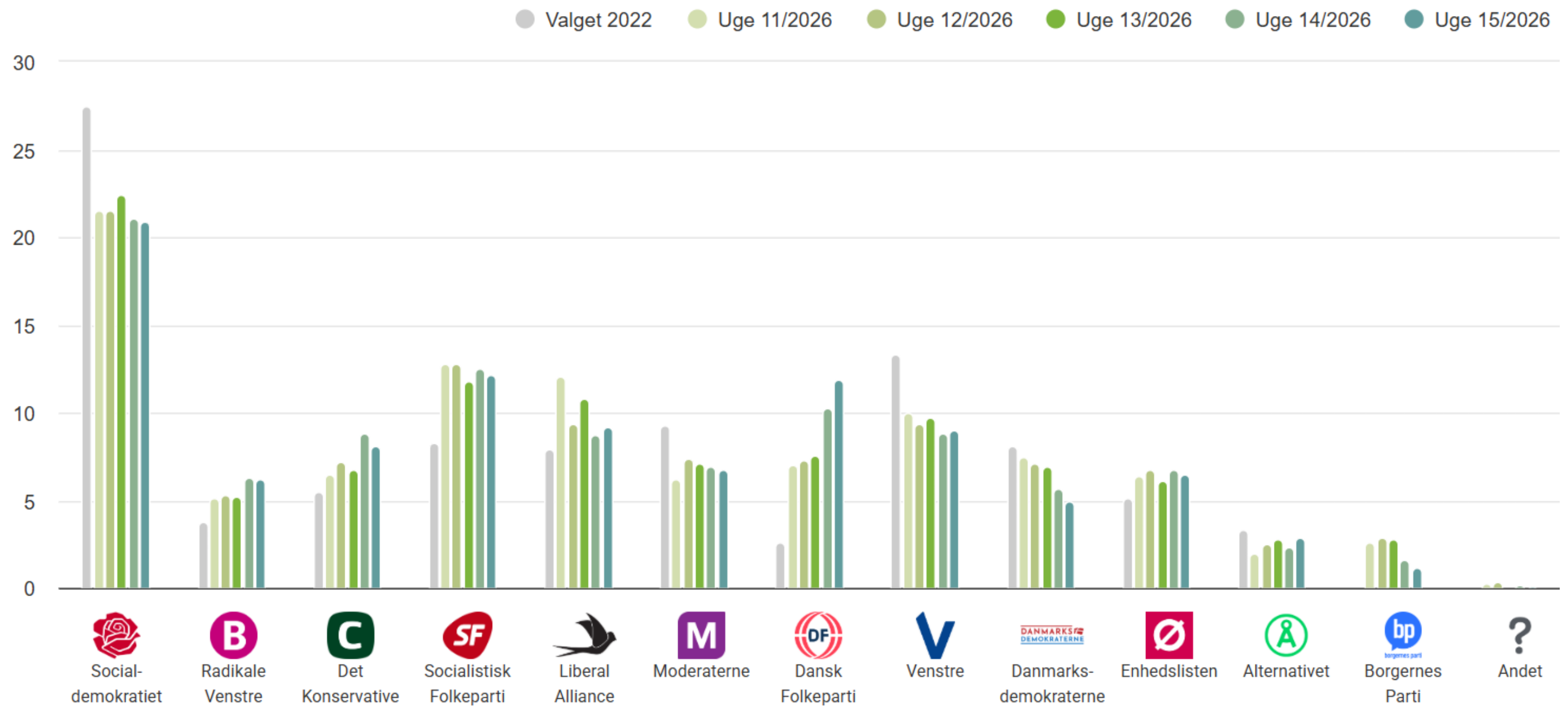


Binomialtest

Begrebsafklaring

- **Population:** En mængde af ensartede individer, objekter eller enheder.
- **Stikprøve:** En lille delmængde af en population. Stikprøverne skal udgøre et repræsentativt udsnit af populationen. Tilfældig udvælgelse.



Stikprøve fra tre populationer

Population	Stikprøve
Den danske vælgerskare	1000 vælgere fra Danmark
Alle agurker produceret i et bestemt væksthus	1768 agurker fra væksthuset
Alle mobiltelefoner	1700 mobiltelefoner

Systematiske fejl

- Hvis den tilfældige udvælgelse ikke er så tilfældig som man tror.
- Nogle grupper bliver under eller overrepræsenteret.

Lad os se på et par eksempler…

Systematiske fejl

Undersøgelse	Systematisk fejl
Et kosmetikfirma vil lave en undersøgelse af en bestemt type håndcreme blandt brugerne. Spørgsmålene stilles på firmaets hjemmeside, og tilfældige besøgende på denne side kan svare på dem.	.

Systematiske fejl

Undersøgelse	Systematisk fejl
Et kosmetikfirma vil lave en undersøgelse af en bestemt type håndcreme blandt brugerne. Spørgsmålene stilles på firmaets hjemmeside, og tilfældige besøgende på denne side kan svare på dem.	Kun personer, der besøger hjemmesiden, har mulighed for at deltage. Man kunne forestille sig, at disse er specielt glade for firmaets produkter og derfor ikke er repræsentative for alle brugere. Et andet problem er, at det kun er dem, der har tid og lyst, som deltager i undersøgelsen. Når man selv kan bestemme, om man vil deltage, overrepræsenteres en bestemt persontype.

Systematiske fejl

Undersøgelse	Systematisk fejl
Nogle elever vil i en by undersøge holdningen til en bestemt type kriminalitet. De stiller sig derfor op på byens strøg en onsdag formiddag og udvælger tilfældige forbipasserende til deres stikprøve.	

Systematiske fejl

Undersøgelse	Systematisk fejl
Nogle elever vil i en by undersøge holdningen til en bestemt type kriminalitet. De stiller sig derfor op på byens strøg en onsdag formiddag og udvælger tilfældige forbipasserende til deres stikprøve.	Personer, der går på byens strøg en onsdag formiddag, er ikke repræsentative for alle byens borgere. Fx vil mange mennesker med fast arbejde være underrepræsenteret.

Systematiske fejl

Undersøgelse	Systematisk fejl
Et meningsmålingsinstitut ringer 1000 tilfældigt genererede telefonnumre op. Hvis opkaldet besvares, og personen er mindst 18 år, spørger man om, hvad vedkommende ville stemme, hvis der var valg i morgen. I modsat fald vælges et andet tilfældigt telefonnummer.	

Systematiske fejl

Undersøgelse	Systematisk fejl
Et meningsmålingsinstitut ringer 1000 tilfældigt genererede telefonnumre op. Hvis opkaldet besvares, og personen er mindst 18 år, spørger man om, hvad vedkommende ville stemme, hvis der var valg i morgen. I modsat fald vælges et andet tilfældigt telefonnummer.	Man undgår de personer, der ikke svarer på opkald fra "ukendt nummer". Man kan også forestille sig, at personer, der arbejder i weekender og aftentimer bliver underrepræsenteret, da opkaldene ofte foretages på netop disse tidspunkter, hvor flest mennesker kan besvare opkaldet.

Skjulte variable

- ✓ Pas på med forhastede konklusioner.
- ✓ Man kan statistisk vise, at der er en sammenhæng mellem hvor mange stjerneskud der er observeret, og hvor mange, der bliver forkølede.
- ✓ Heraf kan man ikke slutte, at stjerneskud gør folk forkølede. Årstiden er en skjult variabel.
- ✓ Om vinteren kan man se flere stjerneskud end resten af året, og om vinteren er der flere der bliver forkølede end resten af året.
- ✓ De to variable er altså begge påvirket af den samme skjulte variabel.

Eksempler

- https://www.tylervigen.com/spurious-correlations?fbclid=IwAR2W77f25cp7ISQI4Ltk-RsqRNO2P4nERpkCQwXTdKRHLwKaJ1yjhXH7dAA

Hvad er et statistisk test?

- I en del af statistikken er formålet at konkludere noget generelt om en hel population på baggrund af en stikprøve.
- Som et eksempel herpå ser vi nu på binomialtestet, hvor andelen af "succeser" i en stikprøve bruges til at forkaste eller acceptere en hypotese om den tilsvarende andel i hele populationen.

Alternativet
Enhedslisten
Socialistisk Folkeparti
Socialdemokratiet
Radikale Venstre
Venstre
Konservative folkeparti
Nye Borgerlige
Liberal Alliance
Dansk Folkeparti

Eksempel: Er terningen ærlig?

Vi ønsker at bestemme om en terning er ærlig til en julefrokost, hvor det gælder om at slå seksere.

Vi indfører en stokastisk variabel, som tæller antallet af seksere på 100 kast med en terning. En terning påstås normalt at have sandsynligheden $p = \frac{1}{6}$ for at vise seks øjne.

Med antagelsen om at terningen er ærlig, må X være binomialfordelt med:

$$X \sim b\left(100, \frac{1}{6}\right)$$

Fordi:

- Der er 100 uafhængige gentagelser af basiseksperimentet.
- Der er kun to udfald: Enten slå en sekser (succes) eller også gør der ikke (fiasko).
- Sandsynligheden for succes er den samme ved hvert basiseksperiment.

Nulhypotese

- Først opstilles en såkaldt nulhypotese H_0 som er den hypotese, vi stoler på indtil videre. Vi formulerer også en alternativ hypotese H_a, der udtrykker alle andre muligheder end nulhypotesen:

H_0: Sandsynligheden for at slå en sekser er $\frac{1}{6}$ (eller blot $p = \frac{1}{6}$)

H_a: Sandsynligheden for at slå en sekser er <u>ikke</u> $\frac{1}{6}$ (eller blot $p \neq \frac{1}{6}$)

I konklusionen til sidst, vil vi enten acceptere H_0 eller forkaste den til fordel for H_a. Testen vi vil udføre kaldes en dobbeltsidet test, fordi den alternative hypotese både indeholder muligheden $p < \frac{1}{6}$ eller $p > \frac{1}{6}$.

Population, stikprøve og teststørrelse

- **Populationen**, som vi har en hypotese om i dette eksempel, er lidt abstrakt, nemlig "alle kast med terningen".
- Fra den udtages en **tilfældig stikprøve**, dvs. en række kast. Vi vælger at kaste terningen 100 gange og forestiller os, at der er 8 seksere.
- Konklusion på vores test skal konkluderes på dette tal, som kaldes teststørrelsen:

Teststørrelsen = 8

- Bemærk, at vi ville forvente at få følgende antal seksere ved 100 forsøg:

$$\mu = 100 \cdot \frac{1}{6} = 16{,}66666$$

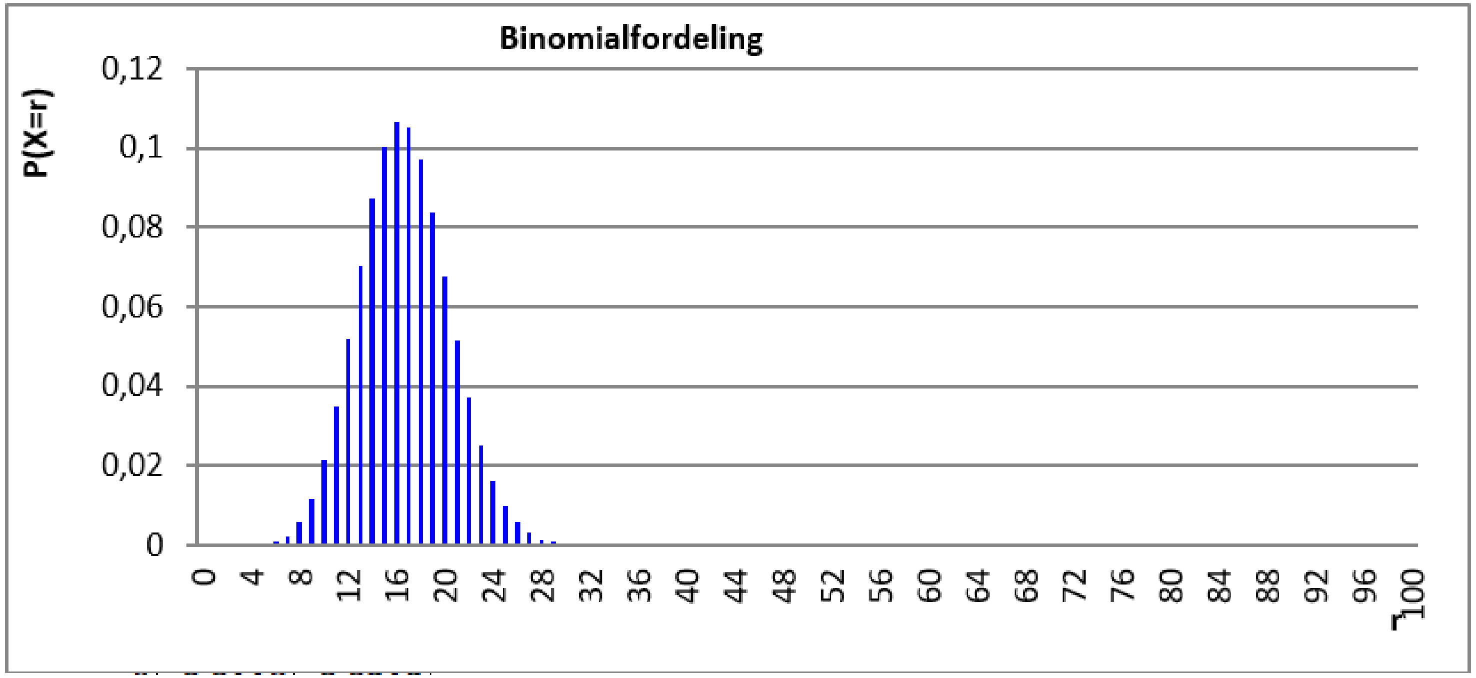
Den forventede værdi = middelværdien

- MEN er det fordi vi har været uheldige med vores stikprøve eller det udtryk for at terningen er uærlig?? Bliver I meget kritiske overfor nulhypotesen, når teststørrelsen er 8?

Acceptmængde, kritisk mængde og signifikansniveau

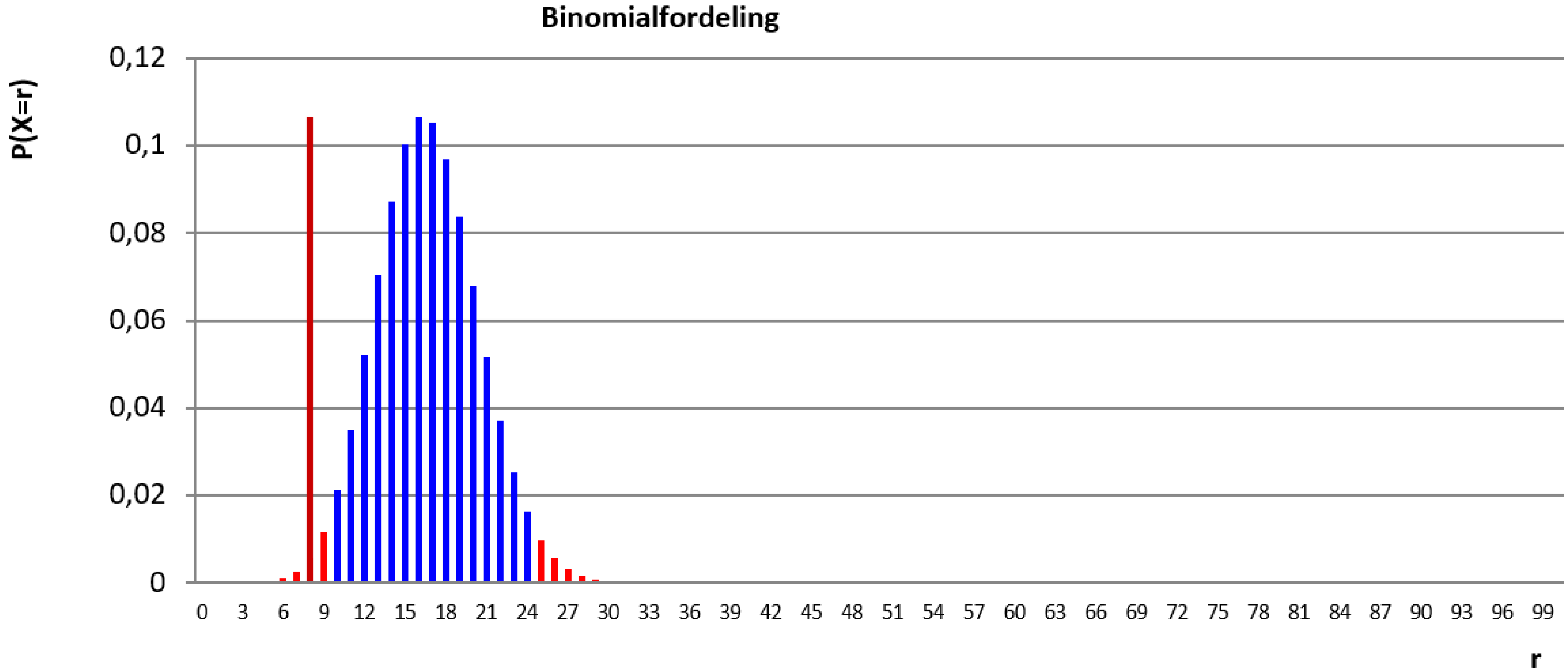
Lad os sige, at vi endnu ikke kender teststørrelsen og generelt betragter 100 kast med en terning, hvor H_0 er en sand hypotese. Da er antallet af seksere en stokastisk variabel (X), der er binomialfordelt $b\left(100, \frac{1}{6}\right)$ med følgende sandsynlighedsfordeling:

Binomialfordeling		P(X=		)=
n=	**100**	P(X≤		)=
p=	**0,1667**	P(	≤X≤	)=



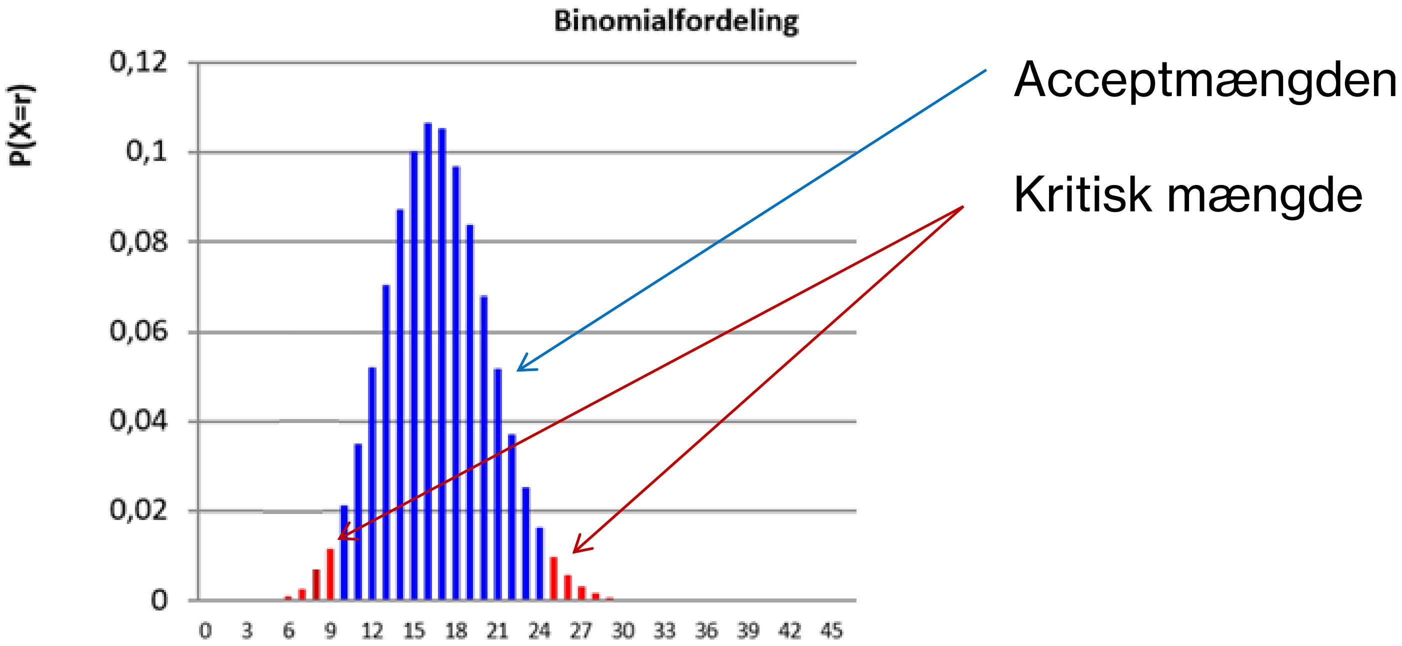
Acceptmængde, kritisk mængde og signifikansniveau

Der er lille sandsynlighed for, at antallet af seksere i en stikprøve afviger meget fra middelværdien (μ = 16,6666). Med dette i mente indfører vi en såkaldt kritisk mængde, K, med en så lille sandsynlighed, at vi simpelthen ikke tror på H_0, hvis teststørrelsen lander her.



Acceptmængde, kritisk mængde og signifikansniveau

På figuren er K markeret med rødt, og da testen er dobbeltsidet består den af to sider, der ligger ca. lige langt fra middelværdien. Sandsynligheden for at havne i K (under forudsætning af, at H0 er sand) kaldes signifikansniveauet og er et tal, vi selv vælger. Ofte vælges et signifikansniveau på 5%, hvilket også er tilfældet her. De to sider i K deles ligeligt om de 5%, forstået på den måde, at de hver især skal have en sandsynlighed, der højst er 2,5% og så tæt på dette tal som muligt.



Acceptmængde, kritisk mængde og signifikansniveau

Hvis vi ønsker at bestemme grænserne for venstre og højre side i K, skal vi se på sandsynligheder af typerne $P(X \leq r)$ og $P(X \geq r)$.

Nedre grænse:

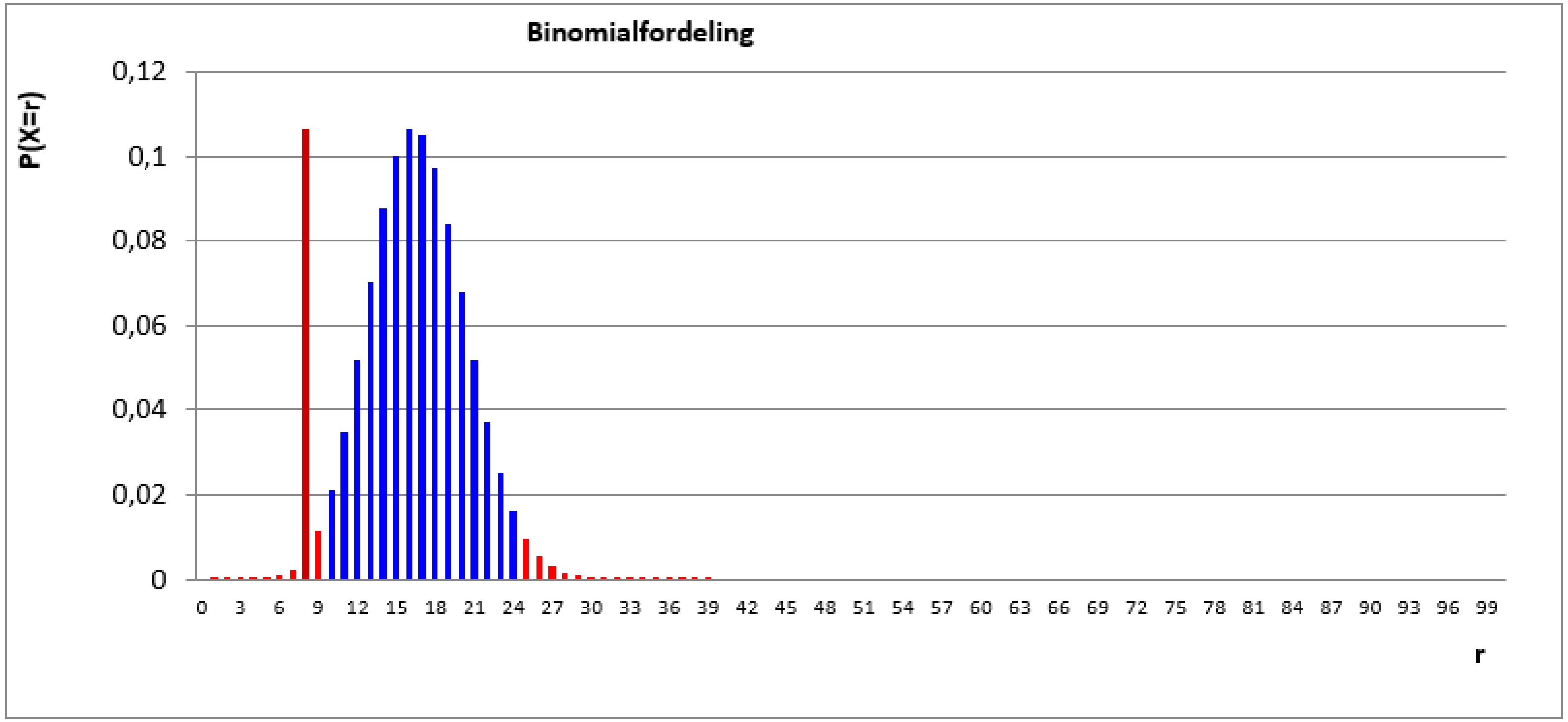
Binomialfordeling			P(X=		)=	
	n=	**100**	P(X≤	9	)=	**0,0213**
	p=	**0,1667**	P(	≤X≤	)=	

Øvre grænse:

Binomialfordeling			P(X=			)=	
	n=	**100**	P(X≤			)=	
	p=	**0,1667**	P(	25 ≤X≤	100	)=	**0,021703**

WordMat = Nemt!

- I WordMat er der en indbygget en Excel-skabelon, som gør det nemt at bestemme acceptmængden og den kritiske mængde.



Kritiske værdier

$k_1 = 9$ $k_2 = 25$

Kritisk mængde: {0,1,...9} U {25,26,...,100}

Den kritiske mængde fremkommer helt automatisk

Konklusion

På baggrund af Excel-skabelonen kan vi altså fastslå, at den kritiske mængde er:

$$K = \{0, 1, 2, \dots, 7, 8, 9, 25, 26, 27, \dots, 100\}$$

Og at acceptmængden er:

$$A = \{10, 11, 12, \dots, 21, 22, 23, 24\}$$

Konklusion:

Altså forkastes nulhypotesen, da teststørrelsen på 8 ligger i den kritiske mængde K.
Altså har vi påvist, på et 5 % signifikansniveau, at terningen er *uærlig*.

Konklusion (alternativ version)

Vi behøver ikke at bestemme de præcise grænser for K for at finde en konklusion. Vi kan nemlig hurtigt undersøge, om teststørrelsen 8 ligger i den kritiske mængde, og i så fald i venstre side, ved at beregne $P(X \leq 8)$ i binomialfordelingen med $b \sim \left(100, \frac{1}{6}\right)$.

Binomialfordeling			P(X=			)=	
	n=	**100**	P(X≤	8		)=	**0,0095**
	p=	**0,1667**	P(		≤X≤		)=

Denne sandsynlighed kaldes for *p*-værdien. Da sandsynligheden ikke overstiger 2,5% (halvdelen af signifikansniveauet), ligger teststørrelsen i K, og vores konklusion er, at nulhypotesen forkastes på signifikansniveau 5%.

Vi vælger altså at tro på den alternative hypotese, svarende til at terningen er uærlig.