Diskrete fordelinger

Fordeling 1. *Binomisk fordeling*
Vi gjentar et stokastisk forsøk n ganger, og lar den stokastiske variabelen X være antall ganger en bestemt hendelse S forekommer. Vi kaller dette en binomisk forsøksrekke om sannsynligheten $p = p(S)$ for at S forekkommer er den samme i hvert av forsøkene, og om forsøkene er uavhengige av hverandre. I så fall kalles X en *binomisk fordelt* stokastisk variabel med parametre n og p. Mulige verdier for X er $X(S) = \{0, 1, 2, \dots, n\}$. Det er vanlig å skrive $q = 1 - p$ for sannsynligheten for at S ikke forekommer.

a) $p(X = i) = \binom{n}{i} p^i q^{n-i}$ for $i = 0,1,\dots,n$ b) $E(X) = np$ c) $\mathrm{Var}(X) = npq$

Typisk eksempel: En flervalgseksamen har 15 spørsmål, og vi velger et tilfeldig svaralternativ blant fire mulige på hvert spørsmål. Hvis vi velger svarene på spørsmålene uavhengige av hverandre, så er antall riktige svar en binomisk fordelt stokastisk variabel med parametre $n = 15$ og $p = 1/4$. Den har derfor forventning $15/4 = 3.75$ og varians $45/16 = 2.8125$.

Resultat: En binomisk fordelt variabel X med parametre n og p kan tilnærmes ved å bruke en normalfordelt variabel med $\mu = np$ og $\sigma^2 = npq = np(1-p)$. De Moivres formel (med heltallskorreksjon) gir da

$$p(X \le x) \approx G(z) \quad \text{med } z = (x + 0.5 - \mu)/\sigma = (x + 0.5 - np)/\sqrt{npq}$$

Fordeling 2. *Geometrisk fordeling*
Vi gjentar et stokastisk forsøk ganger helt til hendelsen S forekommer, og lar den stokastiske variabelen X være antall ganger vi må gjenta forsøket (inkludert forsøket der S forekommer). Dersom forsøksrekken er binomisk og sannsynligheten for at S forekommer i hvert av forsøkene er p, så kalles X en *geometrisk fordelt* stokastisk variabel med parameter p. Mulige verdier for X er $X(S) = \{1, 2, \dots\}$. Det er vanlig å skrive $q = 1 - p$ for sannsynligheten for at S ikke forekommer.

a) $p(X = i) = p\,q^{i-1}$ for $i = 1,\dots$ b) $E(X) = 1/p$ c) $\mathrm{Var}(X) = 1/p^2 - 1/p$

Typisk eksempel: Vi kaster en fair mynt helt til vi får kron første gang. Da er antall kast (inkludert kastet som gir kron) en geometrisk fordelt stokastisk variabel med parameter $p = 1/2$. Den har derfor forventning 2 og varians 2.

Fordeling 3. *Hypergeometrisk fordeling*
En populasjon består av N objekter, hvorav M har en bestemt egenskap. Vi trekker n tilfeldige objekter fra populasjonen uten tilbakelegging, og lar X være antall objekter med den spesielle egenskapen blant de vi har trukket. Da kalles X en *hypergeometrisk* stokastisk variabel med parametre N, M, n. Vi skriver $p = M/N$ for andelen som har den spesielle egenskapen, og $q = 1 - p$. Mulige verdier for X er $X(S) = \{0, 1, \dots, n\}$.

a) $p(X = i) = \dfrac{\binom{M}{i} \cdot \binom{N-M}{n-i}}{\binom{N}{n}}$ for $i = 0,1,\dots,n$ b) $E(X) = np$ c) $\mathrm{Var}(X) = npq\,\dfrac{N-n}{N-1}$

Typisk eksempel: Et vareutvalg innehold $N = 100$ enheter, hvorav $M = 20$ er defekte. Vi kontrollerer et tilfeldig trukket utvalg på $n = 30$ enheter. Antall defekte enheter i utvalget er en hypergeometrisk fordelt stokastisk variabel med paramtre $N = 100$, $M = 20$ og $n = 30$. Den har forventning 6 og standardavvik $4.8 \cdot 70/99 \approx 3.4$.

Fordeling 4. *Poissonfordeling*

La X være antall forekomster av en bestemt hendelse i løpet av et bestemt tidsintervall. Anta at forventet antall hendelser per tidsintervall er konstant lik λ, at hendelsene i disjunkte tidsintervall er uavhengige av hverandre, og at to forekomster av hendelsen ikke kan inntreffe samtidig. Da kalles X en *Poissonfordelt* stokastisk variabel med parameter λ. Mulige verdier for X er $X(S) = \{0, 1, 2, \dots\}$.

a) $p(X = i) = \dfrac{\lambda^i}{i!} \cdot e^{-\lambda}$ for $i = 0, 1, \dots$ b) $E(X) = \lambda$ c) $\mathrm{Var}(X) = \lambda$

Typisk eksempel: En havn registrerer hvor mange skip som ankommer havnen hver dag. I gjennomsnitt ankommer 12 skip er dag. Antall ankomster en tilfeldig valgt dag er en Poissonfordelt stokastisk variabel med parameter $\lambda = 12$. Den har forventning 12 og varians 12.

Resultat: Dersom n er stor og p er liten, så er en binomisk fordelt variabel med parametre n og p tilnærmet Poissonfordelt med parameter $\lambda = np$.

Kontinuerlige fordelinger

Fordeling 5. *Uniform fordeling*

La $a < b$. En kontinuerlig stokastisk variabel X er *uniformt fordelt* på intervallet $[a,b]$ dersom $f(x)$ er en konstant funksjon på dette intervallet, og $f(x) = 0$ ellers. Mulige verdier for X er $X(S) = [a,b]$.

a) $p(i \le X \le j) = \dfrac{j - i}{b - a}$ b) $E(X) = \frac{1}{2}(a + b)$ c) $\mathrm{Var}(X) = \frac{1}{12}(b - a)^2$

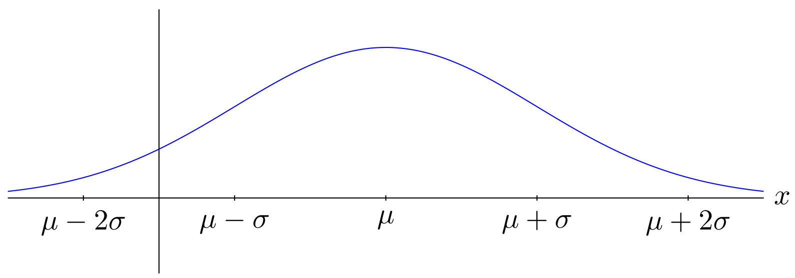
Fordeling 6. *Normalfordeling*

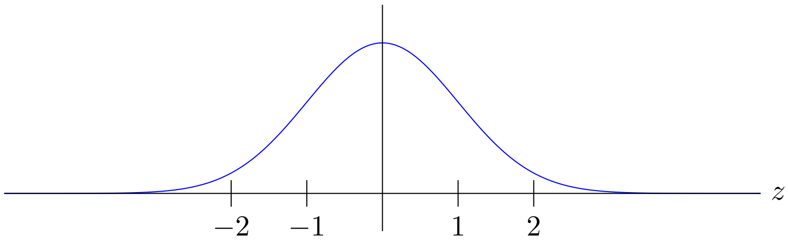
La μ, σ være gitte tall med $\sigma > 0$. Da er en kontinuerlig stokastisk variabel X *normalfordelt* med parametre μ, σ dersom tetthetsfunksjonen til X er gitt ved

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{1}{2}(x-\mu)^2/\sigma^2}$$

Vi skriver noen ganger $N(\mu,\sigma)$ for denne fordelingen. Mulige verdier for X er $X(S) = (-\infty,\infty)$. Tetthetsfunksjonen til normalfordelingen er vist nedenfor, og kalles en Gauss-kurve.

a) $p(a \le X \le b) = \int_a^b f(x)\,\mathrm{d}x$ b) $E(X) = \mu$ c) $\mathrm{Var}(X) = \sigma^2$





Fordeling 7. *Standard-normalfordeling*
Hvis Z har fordeling $N(0{,}1)$, det vil si at Z er normalfordelt med $\mu = 0$ og $\sigma = 1$, kalles Z *standard-normalfordelt*. Hvis X er normalfordelt med fordeling $N(\mu,\sigma)$, så er $Z = (X - \mu)/\sigma$ standard-normalfordelt. Tetthetsfunksjonen til Z med fordeling $N(0{,}1)$ er vist ovenfor. Merk at z-verdiene svarer til antall standardavvik fra $\mu = 0$. Den kumulative fordelingsfunksjonen til Z skrives ofte $G(z) = p(Z \leq z)$, hvor G står for Gauss, og kan beregnes ved hjelp av kalkulator.

a) $p(a \leq Z \leq b) = G(b) - G(a)$ b) $E(Z) = 0$ c) $\mathrm{Var}(Z) = 1$

Fordeling 8. *Eksponentialfordeling*
La X være tiden det tar før første forekomst i en Poisson-prosess med forventet antall forekomster λ per tidsenhet. Da er X *eksponentialfodelt* med parameter λ, og har tetthetsfunksjon

$$f(x) = \lambda\, e^{-\lambda x}, \quad x \geq 0$$

Mulige verdier for X er $X(S) = [0, \infty)$.

a) $p(i \leq X \leq j) = e^{-\lambda i} - e^{-\lambda j}$ b) $E(X) = 1/\lambda$ c) $\mathrm{Var}(X) = 1/\lambda^2$