1 Föreläsning V; Kontinuerlig förd.

Ufallsrummet har hittills varit dsikret, den stokastisk variabeln har endast kunnat anta ett *antal värden*. Ex.vis Poissonfördeln. är antal observationer inom ett tidsintervall men *när*, d.v.s. i vilket tidsintervall en observation görs Ger fördelningen inte svar på.
En stok. var. som kan anta alla värden i ett intervall kallas *kontinuerlig*.

Ex 1 Låt ξ var den tid man väntar tills nästa buss kommer. Utan att veta tidtabellen och med tiominutersintervall mellan bussturerna får vi en Fördelning som är likformig i någon mening. Låt ξ =väntetid på nästa buss. Att behöva vänta högst tiden $t \in [0, 10]$ är

$$P(\xi \leq t) = \frac{t}{10}, \quad 0 \leq t \leq 10.$$

En sådan fördelning kallas *Rektangelfördelning*. Funktionen $P(\xi \leq t) =: F(t)$ kallas *fördelningsfunktion*. Dess derivata $F'(t) = \frac{1}{10} =: f(t)$ kallas *frekvensfunktion* (sannolikhetsfunktion). Man kan åskådliggöra $f(t)$ och $F(t)$ i ett koordinatsystem. Observera att arean mellan $y = f(t)$, $y = 0$, där $0 \leq t \leq 10$ är 1.

∎

Definition 1 För en kontinuerlig stokastisk variabel definierad i ett intervall $I = (a, b)$ är $f(x)$ en frekvensfunktion, om

$$\int_I f(x)dx = 1, \text{ och } f(x) \geq 0. \tag{1}$$

Motsvarande fördelningsfunktion $F(x)$ är

$$\int_a^x f(t)dt = F(x), \quad a < x < b. \tag{2}$$

Kommentarer

- Frekvensfunktionen $f(x)$ ovan motsvarar $P(\xi = x)$ för diskret fördelning. För kontinuerlig fördelning är $P(\xi = x) = 0$, så att frekvensfunktion för diskret och kont.inuerlig fördelning skiljer sig åt.
- Fördelningsfunktion för diskret fördelning är

$$F(x) := P(\xi \leq x) = \sum_{x' \leq x} P(\xi = x').$$

Ex.vis för $\xi \in \mathrm{Geo}(p)$ är

$$F(x) = \sum_{x'=1}^{x} (1-p)^{x'-1} p.$$

- För en kontinuerlig fördelning är $P(\xi < x) = P(\xi \leq x)$, d.v.s. utfallet i en enstaka punkt $P(\xi = x) = 0$.

Ex 2 Betrakta rektangelfördelningen i föreg- exempel. Hur definierar man ett väntevärde för ξ? Det borde vara 5 (min) eftersom det är mitten på intervaller $[0, 10]$. Ett sätt att räkna ut det är att integrera

$$\int_0^{10} tf(t)dt = \frac{1}{10}\left[\frac{t^2}{2}\right]_0^{10} = 5 = E(\xi) = \mu$$

Precis som vi gissade.

■

Kommentarer

- I princip visar exemplet ovan på hur man beräknar ett väntevärde för en godtycklig kontinuerlig fördelning.

1.1 Exponentialfördelning

För denna fördelning är frekvensfunktionen beroende av en parameter $\lambda > 0$ och är

$$f(t) = \begin{cases} 0 & \text{om } t < 0 \\ \lambda e^{-\lambda t} & \text{om } t \geq 0 \end{cases}$$

Motsvarande fördelningsfunktion är

$$F(x) = \begin{cases} 0 & \text{om } x < 0 \\ \displaystyle\int_0^x \lambda e^{-\lambda t} = 1 - e^{-\lambda x} & \text{om } x \geq 0 \end{cases}$$

Valet av t som oberoende variabel är att de ger sannolikhet för livslängd, alltså tid. Vi observerar att

$$\lim_{x \longrightarrow \infty} F(x) = 1.$$

Ex 3 Vi har beräknat för $\xi \in \text{Po}(5.5\,(h))$ sannolikheten för ankommande långtradare under ett dygn. Ex.vis är sannolikheten för *minst* $\xi = 2$ långtradare

$$P(\xi \geq 2) = 1 - P(\xi < 2) = 1 - e^{-5.5}\left(\frac{5.5^0}{0!} + \frac{5.5}{1!}\right).$$

$P(\xi \geq 2) = 1 - P(\xi \leq 1) = 1 - F(1)$, där $F(x)$ är fördelningsfunktionen för ξ.

■

Ex 4 Ett relä i en elektrisk krets har livslängd som är exponentialfördelad med $\lambda = 0.004$ (h^{-1}).

(a) Beräkna sannolikheten att livslängden överskrider 150 h.

(b) Beräkna sannolikheten att livsländen överskrider 200 h, om livslängden har överskridit 50 h.

Lösning:

(a) Sannolikheten att livslängden överskrider 150 h är

$$1 - F(150) = 1 - (1 - e^{-0.004\cdot 150}) = 0.548812...$$

(b) Sannolikheten att livsländen överskrider 200 h, om livslängden har överskridit 50 h är en betingad sannolikhet. Sättt A =händelsen att reläts livslängd är ≥ 50 h och B =händelsen att reläts livslängd är ≥ 150 h. Vi skal alltså beräkna

$$P(B|A) = \frac{e^{-\lambda\cdot 200}}{e^{-\lambda\cdot 50}} = e^{-\lambda\cdot 150} = 0.548812...$$

Alltså samma sannolikhet som i (a).

■

Kommentarer

- Vi ser att sannolikheten för att livslängden är ytterligare 150 h om livslängden har överskridit 50 h är densamma som för att livslängden är 150 h från det att relät är nytt. Det betyder att relä inte blir sämre (eller bättre) med tiden. Har det överlevt ett tidsintervall $[0, t_0]$ är sannolikheten att den överlever t.o.m. $t_0 + t_1$, densamma som att den överlever tidsintervallet $[0, t_1]$.
- Exponentialfördelning beskriver sannolikhet för icke-åldrande och är ett specialfall av Weibullfördelning. Den senare beskriver olika typer av åldrande, ex.vis elektriska eller andra komponenter som blir bättre till en viss tidpunkt och sedan åldras/blir sämre.
- Förväntad överlevnadstid är (partiell integration behövs)

$$E(\xi) = \int_0^\infty t \cdot e^{-\lambda t} dt = \frac{1}{\lambda}.$$

- I exemplet ovan är förväntad livslängd således

$$\frac{1}{0.004} = 250 \text{ h}.$$

- För radioaktivitet används exponentialfördelning; Livslängden för an individuell atoms radioaktivitet $\in \text{Exp}(\lambda)$ för ett visst λ beroende på vilken isotop det rör sig om.

1.2 Normalfördelning

Frekvensfunktion är

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

Den har två parametrar μ som är dess väntvärde och σ som är dess standardavvikelse. Man säger att $\xi \in \text{N}(\mu, \sigma)$. Det är komplicerat att beräkna fördelningsfunktionens värde för olika x, d.v.s. att beräkna

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{(t-\mu)^2/(2\sigma^2)} dt. \tag{3}$$

Det är en utmaning att visa att $F(x) \longrightarrow 1$, då $x \longrightarrow \infty$ men man visar i en kurs i *Flervariabelanalys* att

$$\int_{\infty}^{\infty} e^{-t^2/t} dt = \sqrt{2\pi}.$$

Först skall vi göra en variabelsubst. till standardnormalfördelningen , som har väntevärde 0 och standardavvikelse 1. Denna har i sin tur motsvarande fördelningsfunkton $\Phi(x)$ tabellerad (se tabellsamlingen).

1.2.1 V.S. i normalfördelning

Vi gör nu en V.S. från en allm. Normalförd. till standardnormalfördelningen. $\phi(x)$ står för motsvarande frekvensfunktion och $\Phi(x)$ för dito fördelningsfunktion. Sätt i (3)

$$y = \frac{t-\mu}{\sigma} \text{ dom ger } dt = \sigma\, dy$$

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{(x-\mu)/\sigma} e^{-y^2/2} \sigma\, dy = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

Ex 5 En elkabels diameter i cm är normalfördelad $N(0.8, 0.02)$. Vad är sannolikheten att diametern är

(a) högst 0.82 cm,

(b) mer än 0.81 cm,

(c) mellan 0.77 cm och 0.83 cm?

Lösning:

(a) Sannolikheten för högst 0.82 cm är

$$\Phi\left(\frac{0.82-0.80}{0.02}\right) = \{\text{tabell}\} = 0.841345... \approx 0.84$$

(b) Sannolikheten för mer än 0.81 cm är sannolikheten 1 minus sannolikheten för högst 0.81 cm. Den är således

$$1 - P(\xi \leq 0.81) = 1 - \Phi\left(\frac{0.81-0.80}{0.02}\right) = 0.308538 \approx 0.3.$$

(c) Denna sannolikhet, som vi kallar p, får vi som en differens.

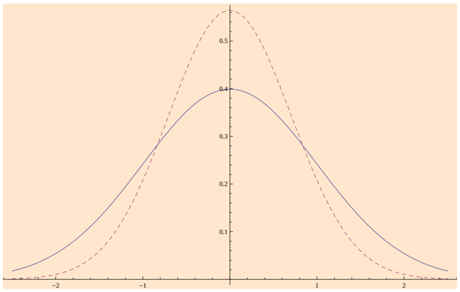
$$\begin{aligned} p &= F(0.83) - F(0.77) = \Phi\left(\frac{0.83-0.80}{0.02}\right) - \Phi\left(\frac{0.77-0.80}{0.02}\right) = \\ &= \Phi(1.5) - \phi(-1.5) = 2\Phi(1.5) - 1 = \{\text{tabell}\} = 0.866386 \approx 0.87\,. \end{aligned}$$

Detta är alltså sannolikheten att en viss elkabel har en diameter mellan 0.77 cm och 0.83 cm.

■

1.3 Jämförelse: σ_1 och σ_2

Vi passar på att jämföra grafen av två normafördelningars frekvensfunktioner med samma väntevärde (= 0) och med olika väntevärden. En som inte är streckad är standardnormafördelningen med $\sigma_1 = 1$ och den med streckad graf måste ha mindre standardavvikelse, i själva verket $\sigma_2 = \frac{1}{\sqrt{2}}$.



Kurvan med stˆrst maximum har minst standardavvikelss ($\sigma_2 = \frac{1}{\sqrt{2}}$).