

Föreläsning 8 för TNIU23

Integraler och statistik

Krzysztof Marciniak
ITN, Campus Norrköping,
krzma@itn.liu.se
www.itn.liu.se/~krzma

ver. 2 - 2009-02-06

1 Inledning - lite om statistik

Statistik är en gren av tillämpad matematik som bygger på *sannolikhets teori* och som uttalar sig om matematiska egenskaper av stora populationer eller långa serier av experiment. Den tillämpas (ibland lite för fritt) av politiker och tjänstemän för att fatta viktiga beslut som rör hela samhället. Den här föreläsningen är dock ingalunda någon inledning i statistik utan den skall enbart illustrera hur integraler och integralkalkyl kan *tillämpas* i statistik. Mycket av materialet presenterad nedan kan hittas (utspridd) t.ex. i [1] därifrån ett antal uppgifter hämtades. En modern presentation av statistik utifrån integralteori (som heter mätteori nuförtiden) kan hittas t.ex. i [2].

2 Kontinuerliga stokastiska variabler

Betrakta en variabel X som vid olika tillfällen (olika mätmoment) kan anta olika värden ur (en delmängd av) reella axeln \mathbf{R} . Anta också att dessa värden kan fås med en viss *sannolikhets* som kan fås experimentellt efter många upprepade mätningar. En sådan variabel kallas då för en *kontinuerlig stokastisk variabel* (s.v.). Notera att det inte är någon stringent definition av s.v. då den sistnämnda kräver kunskaper i sannolikhets teori. Vi får nöja oss alltså med den intuitiva beskrivningen ovan. Några exempel på kontinuerliga s.v. listas nedanför:

- Längd (eller vikt el. dyl.) för en slumpmassigt vald man i landet
- Väntetid på en buss vid en busshållplats
- Överlevnadstid för en radioaktiv atom
- Den årliga inkomsten hos en slumpmassigt vald person i landet. Denna variabel är egentligen diskret med minsta "kvant" = 1 öre men vi kan anta att den är så gott som kontinuerlig.

Notera att i alla dessa exempel den s.v. antar vissa värden mer sannolikt än andra. För att beskriva detta inför man begreppet *täthetsfunktion*.

Definition 1 En (Riemannintegrerbar) funktion $f : \mathbf{R} \rightarrow \mathbf{R}$ som uppfyller två kriterier:

1. $f(x) \geq 0$ för alla x
2.
$$\int_{-\infty}^{\infty} f(x) dx = 1$$
 (1)

kallas för täthetsfunktion för en stokastisk variabel X .

Notera att en täthetsfunktion behöver *ej* vara kontinuerlig. En täthetsfunktion definierar entydigt hur sannolikt det är att den tillhörande s.v. X antar ett värde inom en viss mängd. Man tolkar nämligen uttrycket $f(x)dx$ som den *infinitesimala sannolikheten* att X antar ett värde mellan x och $x + dx$. Som konsekvens får vi då att **sannolikheten att en kontinuerlig stokastisk variabel X med täthetsfunktionen f antar vid ett givet tillfälle värdet i mängden $A \subset \mathbf{R}$ ges av**

$$P(X \in A) = \int_A f(x) dx \quad (2)$$

där alltså symbolen A under integraltecknet betyder att vi integrerar f på mängden A . Mängden A kan (men behöver ej vara) ett intervall eller en union av intervaller (det är i princip det vi kan från analyskursen). Speciellt, sannolikheten $P(a \leq X \leq b)$ för en s.v. X med täthetsfunktionen $f(x)$ att anta värdet någonstans i intervallet $[a, b]$ ges av

$$P(a \leq X \leq b) = \int_a^b f(x) dx \quad (3)$$

Ur detta följer då att

$$P(X = c) = \int_c^c f(x) dx = 0$$

så att sannolikheten $P(X = c)$ att X får exakt värdet c är noll! Således, sannolikheterna $P(a \leq X \leq b)$, $P(a < X < b)$, $P(a \leq X < b)$ och $P(a < X \leq b)$ är alla lika och ges av formeln (3). En annan situation uppstår när f är så kallad *distribution*, men vi skall inte studera detta här (se dock uppgift 6 för att se vad kan hända då). Vidare, notera att

$$P(-\infty < X < \infty) = \int_{-\infty}^{\infty} f(x) dx = 1$$

enligt (1) vilket avspeglar det självklara påståendet att sannolikheten för att en s.v. X antar *något* värde vid ett experiment är lika med 1 och som således är en källa till Definition 1.

Problem 2 Bestäm konstanten c så att funktionen

$$f(x) = \begin{cases} cx^2 & \text{för } 0 < x < 6 \\ 0 & \text{annars} \end{cases}$$

blir en täthetsfunktion. Bestäm sedan sannolikheten att den tillhörande s.v. X antar värdet större än 3. **Lösning:** Enligt (1), vi kräver först att $f(x) \geq 0$ vilket medför att $c > 0$ ($c = 0$ är uteslutet, varför?). Vidare, vi kräver att $\int_{-\infty}^{\infty} f(x) dx = 1$ vilket ger

$$\int_{-\infty}^{\infty} f(x) dx = \int_0^6 cx^2 dx = c \left[\frac{x^3}{3} \right]_0^6 = 72c = 1$$

så att $c = \frac{1}{72}$. Vidare, enligt (3), den sökta sannolikheten blir

$$P(3 < X) = \int_3^{\infty} f(x) dx = \frac{1}{72} \int_3^6 x^2 dx = \frac{1}{72} \left[\frac{x^3}{3} \right]_3^6 = \frac{1}{216} (216 - 27) = \frac{189}{216} = \frac{7}{8} = 0,875$$

Nedanför följer ett antal typiska täthetsfunktioner:

1. Likförmig fördelning: en s.v. X sägs vara likförmigt fördelad om dess täthetsfunktion har formen

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{för } a < x < b \\ 0 & \text{annars} \end{cases} \quad (4)$$

Denna fördelning beskriver en s.v. som kan anta alla värden ur intervallet $]a, b[$ (eller $[a, b]$ etc.) med samma sannolikhet.

2. Exponentialfördelning: ges av täthetsfunktionen

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{för } x > 0 \\ 0 & \text{för } x \leq 0 \end{cases} \quad (5)$$

där $\lambda > 0$.

3. Normalfördelning: ges av

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (6)$$

(med $\sigma > 0$) och används mycket ofta i bl.a. populationsbiologi. Den beskriver en typisk fördelning av en egenskap hos en population. Mer om μ och σ under fö.

Exempel 3 (Väntetiden för en buss). Antag att väntetiden T för en buss är likförmigt fördelad med $a = 0$ och $b = 10$ i (4) (bussar anländer var 10:e minut). Sannolikheten att du skall vänta på bussen mer än p minuter (med $0 \leq p \leq 10$) ges då av

$$P(T > p) = \int_p^{10} \frac{1}{10-0} dt = \frac{1}{10} [t]_p^{10} = \frac{1}{10}(10-p) = 1 - \frac{p}{10}$$

och vid $p = 10$ är noll (bussen kommer säkert efter högst 10 minuter).

Problem 4 (Radioaktivt sönderfall). Exponentialfördelningen (5) beskriver bl.a. radioaktivt sönderfall: sannolikheten att en radioaktiv atom överlever en viss tid T ges av (5). Bestäm tiden $t_{1/2}$ vid vilken hälften av atomer av samma typ försvinner (s.k. halveringstid) som funktion av λ .

Lösning: den sökta tiden uppfyller sambandet

$$P(0 < T < t_{1/2}) = \int_0^{t_{1/2}} \lambda e^{-\lambda t} dt = \frac{1}{2}$$

Å andra sidan

$$\int_0^{t_{1/2}} \lambda e^{-\lambda t} dt = \lambda \left[\frac{e^{-\lambda t}}{-\lambda} \right]_0^{t_{1/2}} = -[e^{-\lambda t}]_0^{t_{1/2}} = 1 - e^{-\lambda t_{1/2}}$$

vilket leder till sambandet

$$e^{-\lambda t_{1/2}} = \frac{1}{2}$$

som ger

$$t_{1/2} = \frac{\ln 2}{\lambda}.$$

Värden för T kan variera från mikrosekunder till miljarder år.

3 Fördelningsfunktion, median, kvantiler och kvar(!)tiler

Av speciellt vikt är sannolikheten att en s.v. X antar värden mindre än eller lika med x . Denna sannolikhet får ett särskilt namn:

Definition 5 Funktionen

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt$$

kallas *fördelningsfunktionen* för den s.v. X .

Således, sannolikheten att den s.v. X antar ett värde ur intervallet mellan a och b ges av

$$P(a < X \leq b) = F(b) - F(a)$$

och är inget annat än insättningsformeln vi läste om! (se analysboken Sats 6.8 sid. 292). Med all rätt kan ni nu även misstänka att följande samband är sant:

Sats 6 Om f är kontinuerlig i punkten x gäller att

$$F'(x) = f(x).$$

Det betyder att **fördelningsfunktionen F är en primitivfunktion för täthetsfunktionen f** -med bägge tillhörande förstås samma s.v. (detta är ju helt enkelt analysens huvudsats, Sats 6.7 sid. 290 i analysboken).

Definitionen 5 och egenskaperna (1) hos varje täthetsfunktion leder till följande allmänna observation:

Sats 7 För varje fördelningsfunktion F gäller att

1. $\lim_{x \rightarrow -\infty} F(x) = 0$, $\lim_{x \rightarrow \infty} F(x) = 1$
2. F är en växande (ej nödvändigtvis strängt) funktion
3. F är högerkontinuerlig för varje x

Notera alltså att F behöver ej vara kontinuerlig utan bara *högerkontinuerlig* i varje punkt. Om F är diskontinuerlig i en punkt så är motsvarande täthetsfunktion en distribution (se också uppgift 6).

Exempel 8 Genom integration visar man lätt att den likformiga fördelningen (4) har följande fördelningsfunktion:

$$F(x) = \int_{-\infty}^x f(t) dt = \begin{cases} 0 & \text{om } x < a \\ \frac{x-a}{b-a} & \text{om } a < x \leq b \\ 1 & \text{om } x > b \end{cases} \quad (7)$$

Se Figur 3.8 sid. 59 i [1].

Exempel 9 Samma procedur visar att exponentialfördelningens (5) fördelningsfunktion är

$$F(x) = \int_{-\infty}^x f(t) dt = \begin{cases} 0 & \text{om } x < 0 \\ 1 - e^{-\lambda x} & \text{om } 0 \leq x \end{cases}$$

Ofta behöver man i statistiska undersökningar svara på följande fråga: givet ett värde α ($\alpha \in [0, 1]$), hur stort x måste vara så att sannolikheten att den s.v. X antar högst värdet x blir exakt lika med α ?

Definition 10 Antag att $\alpha \in [0, 1]$. Lösningen till ekvationen

$$P(X \leq x) = \alpha$$

d.v.s. till ekvationen

$$F(x) = \alpha \quad (8)$$

kallas då α -kvantilen för den s.v. X och betecknas med x_α .

Av speciell vikt är $\frac{1}{2}$ -kvantilen $x_{0,50}$ som heter *median* och som anger det värde på x för vilken sannolikheten att en s.v. X antar värdet $\leq x$ är exakt $1/2$. T.ex: medianlönen i en population är den lönenivån för vilken hälften av populationen tjänar mindre än just medianlönen. Medianlönen beskriver lönespridningen i en population mycket bättre än medellönen för den påverkas inte av att enstaka få personer (t.ex. chefer i ett stort företag) tjänar mycket bättre än den övriga populationen (den övriga personalen).

Definition 11 *Kvantiler $x_{0,25}$, $x_{0,50}$ och $x_{0,75}$ kallas respektive för undre kvaRtilen, medianen respektive övre kvaRtilen.*

Notera att i [1] har man en annan definition av x_α med $1 - \alpha$ istället av a i (8).

Exempel 12 *Medianen för normalfördelningen (6) är lika med μ . Varför?*

Problem 13 *Beräkna medianen samt de övriga kvartilerna för den likformiga fördelningen (4). **Lösning:** det går att gissa förstås att medianen är $\frac{a+b}{2}$. Vill vi göra det formellt löser vi ekvationen (8) med F som i (7) och med $\alpha = \frac{1}{2}$. Detta ger*

$$\frac{x-a}{b-a} = \frac{1}{2}$$

som ger det angivna värdet $x_{0,50} = \frac{a+b}{2}$. Rent allmänt, för att få kvantilen x_α måste vi lösa ekvationen

$$\frac{x-a}{b-a} = \alpha$$

som ger

$$x_\alpha = (1-\alpha)a + \alpha b.$$

(det kan också gissas!). Således: den undre kvartilen är $x_{0,25} = \frac{3}{4}a + \frac{1}{4}b$ medan den övre kvartilen är $x_{0,75} = \frac{1}{4}a + \frac{3}{4}b$.

4 Väntevärdet och varians eller vad vi kan förvänta oss av en stokastisk variabel.

Givet en s.v. X , vad kan vi då säga om det förväntade utfallet av varje mätning av X ? Svaret ges nedan i form av följande definition.

Definition 14 *Väntevärdet $E(X)$ för den s.v. X med täthetsfunktionen $f(x)$ är talet*

$$E(X) = \int_{-\infty}^{\infty} xf(x) dx \tag{9}$$

En annan benämning på talet $E(X)$ är *förväntat värde*.

Problem 15 *Beräkna $E(X)$ för den likformiga fördelningen (4). **Lösning:***

$$E(X) = \int_{-\infty}^{\infty} xf(x) dx = \frac{1}{b-a} \int_a^b x dx = \frac{1}{b-a} \left[\frac{x^2}{2} \right]_a^b = \frac{1}{2}(a+b)$$

Problem 16 *Beräkna $E(X)$ för normalfördelningen (6) med $\mu = 0$. **Lösning:***

$$E(X) = \int_{-\infty}^{\infty} xf(x) dx = E(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} xe^{-\frac{x^2}{2\sigma^2}} dx = \left[\begin{array}{l} \text{variabelbyte:} \\ \frac{x^2}{2\sigma^2} = y \end{array} \right] = \dots = 0$$

vilket lätt ses också genom det att integranden $xf(x)$ är en udda funktion. Rent allmänt, $E(X) = \mu$ för normalfördelningen (6).

Problem 17 Beräkna förväntad överlevnadstid τ för en radioaktiv atom som funktion av λ . **Lösning:** Den förväntade överlevnadstiden ges av

$$\tau = E(T) = \int_{-\infty}^{\infty} tf(t) dt = \int_0^{\infty} t\lambda e^{-\lambda t} dt = [P.I.] = \frac{1}{\lambda},$$

vilket också ger att $\tau = \frac{t_{1/2}}{\ln 2} > t_{1/2}$. En radioaktiv atom förväntas alltså att överleva halveringstiden.

Ofta är vi intresserade i frågan **hur pass utspridda enstaka mätvärden av mätningar av en s.v. kommer att vara**. Denna utspridning mäts med hjälp av talet som karakteriserar den stokastiska variabeln och som kallas *varians*.

Definition 18 Variansen $V(X)$ för en s.v. X med $\mu = E(X)$ är talet som ges av

$$V(X) = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx$$

Problem 19 Beräkna $E(X)$ och $V(X)$ för den s.v. som har täthetsfunktionen

$$f(x) = \begin{cases} 0 & \text{för } x < 1 \\ 3x^{-4} & \text{för } x \geq 1 \end{cases}$$

Lösning:

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} xf(x) dx = E(X) = 3 \int_1^{\infty} x^{-3} dx = 3 \lim_{R \rightarrow \infty} \left[\frac{x^{-2}}{-2} \right]_1^R = -\frac{3}{2} \lim_{R \rightarrow \infty} \left(\frac{1}{R^2} - 1 \right) = \frac{3}{2}. \\ V(X) &= \int_{-\infty}^{\infty} \left(x - \frac{3}{2} \right)^2 f(x) dx = 3 \int_1^{\infty} \left(x - \frac{3}{2} \right)^2 x^{-4} dx = \dots = 3 \left[-\frac{3}{4x^3} - \frac{1}{x} + \frac{3}{2x^2} \right]_1^{\infty} = \frac{3}{4}. \end{aligned}$$

Vi avslutar vår mycket kort resa genom statistik med en intressant sats. Betrakta den s.v. X^2 som alltså är en s.v. kopplad till X på så sätt att varje gång vi mäter X och får resultatet x antar vi att X^2 's värde blev x^2 . Täthetsfunktionen $g(x)$ för X^2 är knuten till täthetsfunktionen $f(x)$ för X via det kanske lite oväntade sambandet

$$g(x) = \begin{cases} 0 & \text{för } x \leq 0 \\ \frac{1}{2\sqrt{x}} (f(\sqrt{x}) + f(-\sqrt{x})) & \text{för } x \geq 0 \end{cases} \quad (10)$$

som bevisas via analysens huvudsats (beviset kan hittas i Appendix och alla studenter förväntas läsa det). Man kan då visa (se Appendix) att

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx \quad (11)$$

vilket tolkas som ett rimligt resultat.

Sats 20 Följande samband gäller

$$V(X) = E(X^2) - [E(X)]^2 \quad (12)$$

där X^2 är en s.v. med täthetsfunktionen g som i (10).

Bevis. Med beteckningar som ovan

$$\begin{aligned} V(X) &= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx = \int_{-\infty}^{\infty} x^2 f(x) dx - 2\mu \int_{-\infty}^{\infty} xf(x) dx + \mu^2 \int_{-\infty}^{\infty} f(x) dx = \\ &= \int_{-\infty}^{\infty} x^2 f(x) dx - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 = E(X^2) - [E(X)]^2 \end{aligned}$$

enligt (11). ■

Problem 21 Beräkna $V(X)$ för den likformiga fördelningen (4) genom att utnyttja sambandet (12).

Lösning: I Exempel 15 fick vi att $E(X) = \frac{a+b}{2}$. Vidare

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x) dx = \frac{1}{b-a} \int_a^b x^2 dx = \frac{b^3 - a^3}{3(b-a)} = \frac{(b-a)(b^2 + ab + a^2)}{3(b-a)} = \\ &= \frac{1}{3}(b^2 + ab + a^2). \end{aligned}$$

Följaktligen

$$V(X) = E(X^2) - [E(X)]^2 = \frac{1}{3}(b^2 + ab + a^2) - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}.$$

5 Appendix

Vi bevisar här formlerna (10) och (11). Vi börjar med (10). Låt oss beteckna X^2 :s fördelningsfunktion och täthetsfunktion med $G(x)$ respektive $g(x)$. Vi har då

$$G(x) = P(X^2 \leq x) = \begin{cases} P(-\sqrt{x} \leq X \leq \sqrt{x}) & \text{om } x \geq 0 \\ 0 & \text{ifall } x < 0 \end{cases}$$

Således, om $x < 0$ har vi att $g(x) = G'(x) = 0$. Återstår att visa formeln (10) för $X \geq 0$. Vi har då enligt formeln ovan samt *analysens huvudsats*

$$\begin{aligned} g(x) &= G'(x) = \frac{d}{dx} P(-\sqrt{x} \leq X \leq \sqrt{x}) = \frac{d}{dx} \int_{-\sqrt{x}}^{\sqrt{x}} f(t) dt = \\ &= f(\sqrt{x}) \cdot \frac{1}{2\sqrt{x}} - f(-\sqrt{x}) \cdot \frac{-1}{2\sqrt{x}} = \frac{1}{2\sqrt{x}} (f(\sqrt{x}) + f(-\sqrt{x})) \end{aligned}$$

v.s.v. Dags att visa (11):

$$E(X^2) = \int_{-\infty}^{\infty} xg(x) dx = \int_0^{\infty} \frac{x}{2\sqrt{x}} (f(\sqrt{x}) + f(-\sqrt{x})) dx = \frac{1}{2} \int_0^{\infty} \sqrt{x} (f(\sqrt{x}) + f(-\sqrt{x})) dx \quad (13)$$

där vi alltså utnyttjad (10). Vidare

$$\int_0^{\infty} \sqrt{x} f(\sqrt{x}) dx = \left[\begin{array}{l} y = \sqrt{x} \\ dx = 2y dy \end{array} \right] = 2 \int_0^{\infty} y^2 f(y) dy$$

På liknande sätt, med hjälp av variabelbytet $y = -\sqrt{x}$, visar vi att

$$\int_0^{\infty} \sqrt{x} f(-\sqrt{x}) dx = 2 \int_{-\infty}^0 y^2 f(y) dy.$$

Insättning av de två sista formlerna i (13) ger (11).

6 Uppgifter

1. Bestäm konstanten c så att funktionen

$$f(x) = \begin{cases} \frac{c}{\sqrt{x+1}} & \text{för } -1 < x < 1 \\ 0 & \text{annars} \end{cases}$$

blir en täthetsfunktion. Beräkna sedan sannolikheten att en s.v. som hör till f antar ett positivt värde.

Svar: $c = \frac{1}{2\sqrt{2}}$. $P(X > 0) = 1 - \frac{1}{\sqrt{2}} \cong 0.293$.

2. Ta fram fördelningsfunktionen för s.v. i uppgift 1. Rita också den!

Svar:

$$F(x) = \begin{cases} 0 & \text{för } x < -1 \\ \frac{1}{\sqrt{2}}\sqrt{x+1} & \\ 1 & \text{för } x \geq 1 \end{cases}$$

3. Visa att funktionen i (5) är verkligen en täthetsfunktion.

Tips: integrera f på $[0, \infty[$ ($f = 0$ för $x < 0$).

4. En s.v. X har fördelningsfunktionen

$$F(x) = \begin{cases} 0 & \text{för } x < 1 \\ 1 - \frac{1}{x^2} & \text{för } x \geq 1 \end{cases}$$

Beräkna fördelningens undre kvartil, median och övre kvartil. Ange även värdet för en godtycklig kvantil x_a .

Svar:

$$x_{0.25} = \frac{2}{\sqrt{3}}, \quad x_{50} = \sqrt{2}, \quad x_{0.75} = 2, \quad x_a = \frac{1}{\sqrt{1-a}}$$

5. Visa att funktionen

$$f(x) = \begin{cases} \frac{3}{2}e^{3x} & \text{för } x < 0 \\ \frac{1}{2}e^{-x} & \text{för } x \geq 0 \end{cases}$$

är en täthetsfunktion för en stokastisk variabel X . Beräkna även de tre kvantilerna för X .

Svar: $-\frac{1}{3} \ln 2$, 0 respektive $\ln 2$.

6. Rita upp fördelningsfunktionen

$$F(x) = \begin{cases} 0 & \text{för } x < 1 \\ \frac{1}{3} + \frac{2}{3}(x-1) & \text{för } 1 \leq x \leq 2 \\ 1 & \text{för } x > 2 \end{cases}$$

Beräkna $P(X \leq 5/3)$, $P(X > 3/2)$, $P(4/3 < X < 5/3)$ och $P(X = 1)$.

Svar: $7/9$, $1/3$, $2/9$ respektive $1/3 = F$:s språnget i 1. Notera att F är ej deriverbar här ty den inte är kontinuerlig i 1. Som nämdes i texten innebär detta att f är ingen vanlig funktion utan en distribution.

7. Väntetiden X (i minuter) från öppningsdag tills dess första kund går in i en affär beskrivs av en s.v. med fördelningsfunktionen:

$$F(x) = \begin{cases} 0 & \text{för } x < 0 \\ 1 - e^{-0.4x} & \text{för } x \geq 0 \end{cases}$$

Beräkna sannolikheten att första kunden dröjer

- a) högst 3 minuter

- b) minst 4 minuter
- c) mellan 3 och 4 minuter
- d) högst 3 eller minst 4 minuter
- e) exakt 2 minuter

Svar: a) $1 - e^{-1.2} = 0,699$ b) $e^{-1.6} = 0,202$ c) $e^{-1.2} - e^{-1.6} = 0,099$ d) $1 - e^{-1.2} + e^{-1.6} = 0,901$
e) 0

8. Beräkna kvantilen x_α för en s.v. X med täthetsfunktionen

$$f(x) = \begin{cases} 0 & \text{för } x < 0 \\ 2xe^{-x^2} & \text{för } x \geq 0 \end{cases}$$

Använd sedan miniräknare för att få de numeriska värdena för $x_{0,50}$, $x_{0,9}$ samt $x_{0,99}$.

Svar: $x_\alpha = \sqrt{-\ln(1 - \alpha)}$, $x_{0,50} = 0,833$, $x_{0,9} = 1,517$, $x_{0,99} = 2,146$.

9. Beräkna $E(X)$ och $V(X)$ om X har täthetsfunktionen

$$f(x) = \begin{cases} \frac{2}{a^2}x & \text{för } 0 \leq x \leq a \\ 0 & \text{annars} \end{cases}$$

Svar: $E(X) = \frac{2}{3}a$. $V(X) = \frac{1}{18}a^2$.

10. Den s.v. X har täthetsfunktionen:

$$f(x) = \begin{cases} 2x & \text{för } 0 \leq x \leq 1 \\ 0 & \text{annars} \end{cases}$$

- a) Beräkna väntevärdet μ och standardavvikelsen σ för X .
- b) Beräkna $P(\mu - \sigma < X < \mu + \sigma)$
- c) Beräkna $P(\mu - 2\sigma < X < \mu + 2\sigma)$

Svar: a) $\mu = \frac{2}{3}$, $\sigma = \frac{1}{\sqrt{18}}$ b) $1 - (\mu - \sigma)^2 = \frac{1}{2} + \frac{2\sqrt{2}}{9}$ c) $6\mu\sigma - 3\sigma^2 = \frac{2\sqrt{2}}{3} - \frac{1}{6}$

References

- [1] G. Blom, J. Enger, G. Englund, J. Grandell och L. Holst, "Sannolighetsteori och statistikteori med tillämpningar", Studentlitteratur 2005.
- [2] P. Billingsley, "Probability and measure", Wiley, 1986