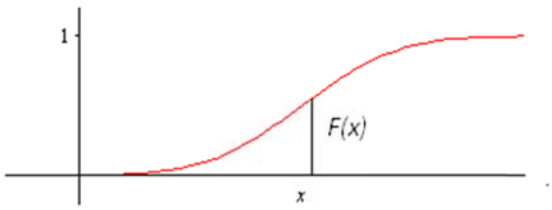
# KONTINUERLIGA STOKASTISKA VARIABLER

( Allmänt om kontinuerliga s.v.)

**Definition 1**. En stokastisk variabel $\xi$. kallas kontinuerlig om **fördelningsfunktionen**

$$F(x) = P(\xi \le x)$$

är kontinuerlig.




**Egenskaper av fördelningsfunktion:**

1) Fördelningsfunktionen $F(x)$ är en **växande** funktion.
2) $0 \le F(x) \le 1$
3) $\lim_{x\to\infty} F(x) = 1$ $, eller\ kortare,$ $F(\infty) = 1$
4) $\lim_{x\to-\infty} F(x) = 0$ $, eller\ kortare,$ $F(-\infty) = 0$

==================================================

Sannolikheten att en kontinuerlig s v har värden i ett intervall (a,b] beräknas enkelt med hjälp av F(x)

$$\boldsymbol{P(a < \xi \le b) = F(b) - F(a)}.$$

Vi bevisar nu att sannolikheten att en kontinuerlig s.v. $\xi$ antar exakt ett givet värde är alltid lika med 0. (Till skillnad från diskreta fördelningar där några punkter bär sannolikhetsmassa)

| |
|---|
| Sats. Om $\xi$ är en kontinuerlig s.v. och b ett reellt tal då är $P(\xi = b) = 0$. |

Låt $a$, $b$ vara reella tal. Eftersom $\{x : \xi = b\} \subseteq \{x : a < \xi \le b\}$ har vi

$$0 \le P(\xi = b) \le P(a < \xi \le b), \text{ dvs}$$

$$0 \le P(\xi = b) \le F(b) - F(a). \quad (*)$$

Eftersom $F(x)$ är **kontinuerlig** gäller $\lim_{a\to b} F(a) = F(b)$. Om vi låter $a \to b$ får vi från (*)

$$0 \le P(\xi = b) \le F(b) - \lim_{a\to b} F(a) \quad (= F(b) - F(b) = 0) \text{ eller}$$

$$0 \le P(\xi = b) \le 0,$$

med andra ord $P(\xi = b) = 0$ **(V.S.B.)**

Följdsatsen:

| |
|---|
| Om $\xi$ är en kontinuerlig s.v. då är $\boldsymbol{P(a < \xi \le b) = P(a \le \xi \le b) = P(a \le \xi < b) = P(a < \xi < b)}$. |

(Vi ser återigen skillnad mellan diskreta och kontinuerliga fördelningar.)

**Definition 2.** Om funktionen $F(x)$ är deriverbar (eventuellt utom i ändligt antal punkter) då kallas derivatan

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

**täthetsfunktionen (** eller **frekvensfunktionen)** för variabeln $\xi$.

**Egenskaper: Täthetsfunktionen(**=frekvensfunktionen) $f(x)$ är en positiv funktion (eftersom F(x) är växande) .

Enligt definitionen gäller $f(x) = F'(x)$ (eventuellt utom i ändligt antal punkter).
Vi antar i vår kurs att $f(x)$ är integrerbar på varje intervall $[a,b]$ . Med sådana antaganden gäller

$$\int_a^b f(x)dx = \int_a^b F'(x)dx = [F(x)]_a^b = F(b) - \mathrm{F(a)} = P(\mathrm{a} < \xi \le b).$$

Därmed

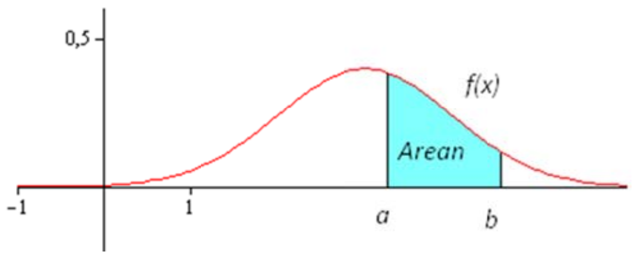
$$\boxed{P(a < \xi \le b) = F(b) - \mathrm{F(a)} = \int_a^b f(x)dx.}$$

Alltså, om vi har $f(x)$ kan vi beräkna sannolikheten $P(a < \xi \le b)$ med hjälp av integralen $\int_a^b f(x)dx$ . Om både $f(x)$ och $F(x)$ är kända så är självklart **enklare** att beräkna sannolikheten $P(a < \xi \le b)$ med hjälp av fördelningsfunktionen $F(x)$, dvs

$$\boxed{P(a < \xi \le b) = F(b) - F(a)}$$

==================================================

**Samband mellan areor under täthetsfunktionen och sannolikheter**




Arean som markeras i ovanstående bild (formeln från analysen) beräknas enligt följande:

**Arean** $= \int_a^b f(x)dx.$

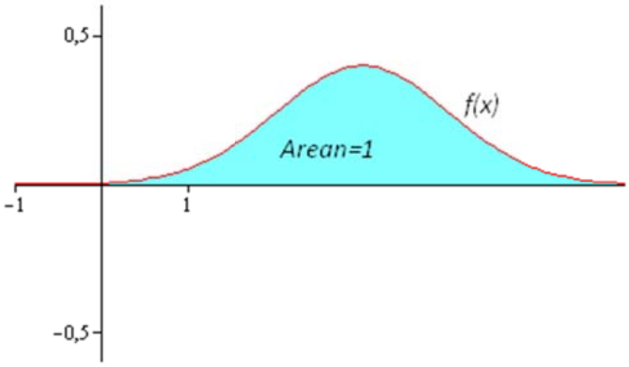
Vi har visat ovan att $\int_a^b f(x)dx = P(a < \xi \le b)$ .

Därför gäller : $P(a < \xi \leq b)$ = arean under täthetsfunktionen ovanpå intervallet [a,b] .

Om vi låter $a$ gå mot $-\infty$ och $b$ gå mot $+\infty$ får vi
$P(-\infty < \xi \leq \infty)$ = arean under hela täthetsfunktionen.

Därmed är arean mellan x-axeln och kurvan $y = f(x)$ lika med 1.

$$P(-\infty < \xi \leq \infty) = \int_{-\infty}^{\infty} f(x)dx = F(\infty) - F(-\infty) = 1 - 0 = 1.$$




===================================================

**Bestämning av $F(x)$ om $f(x)$ är given.**

Vi har visat ovan att $F(x) - F(a) = \int_a^x f(t)dt$. (Vi betecknar övregränsen med $x$ och därför betecknar vi variabeln i integranden med en annan bokstav, $t$.)
Om vi låter $a$ gå mot $-\infty$ får vi ( eftersom $F(-\infty) = 0$ )

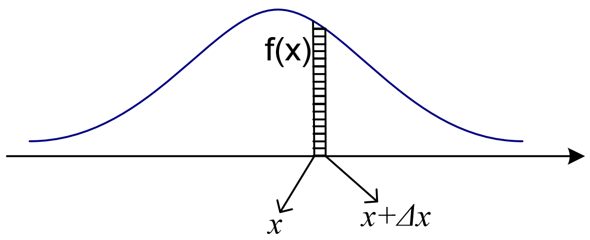
$$F(x) - 0 = \int_{-\infty}^{x} f(t)dt$$

dvs

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

===================================================

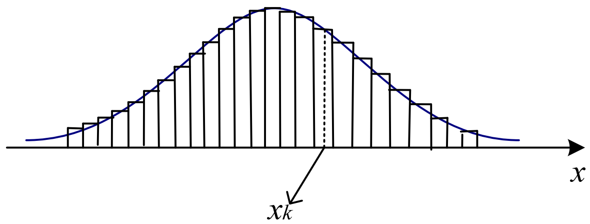
**Approximation av sannolikheten P( x<ξ ≤x+Δx) om Δx är litet:**




Om Δx är litet så är skuggade arean approximativt lika med arean av rektangeln med basen Δx och höjden $f(x)$. Därför kan vi använda följande approximation

$$P(x<\xi \le x+\Delta x) \approx f(x)\Delta x.$$

Vi kan diskretisera en kontinuerlig stokastisk variabel ξ genom att approximera arean under frekvensfunktionen med rektanglar med små baser $\Delta x_k = \Delta x$.




Vi kan betrakta en diskret s.v. X som antar värdena $x_k$ med sannolikheterna $p_k = f(x_k)\,\Delta x$. Då är väntevärdet av den diskreta s.v. X lika med $\sum_k x_k p_k = \sum_k x_k f(x_k)\Delta x$ . Om Δx är litet då är

$\sum_k x_k f(x_k)\Delta x \approx \int_{-\infty}^{\infty} x f(x)dx$ (om integralen existerar). Detta motiverar följande definition.

**VÄNTEVÄRDET** för en kontinuerlig s. v. $\xi$ betecknas m, μ eller $E(\xi)$ och definieras enligt följande

$$\mu = E(\xi) = \int_{-\infty}^{\infty} x f(x)dx$$

På liknande sätt motiveras definitionen av variansen av en kontinuerlig s.v.

**VARIANSEN** av en kontinuerlig s. v. $\xi$ betecknas $V(\xi)$ , Var, $\sigma^2$ eller $s^2$ )

$$V(\xi) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2$$

**STANDARDAVVIKELSEN :** ( Betecknas $\sigma$, $s$ , eller $D(\xi)$ )

$$\sigma = \sqrt{Variansen}$$

**MEDIANEN** definieras som lösningen till ekvationen

$$F(x) = 0.50$$

Medianen delar arean under frekvensfunktionen **(** eller täthetsfunktionen**)** $f(x)$ i två lika delar. Om frekvensfunktionen är symmetrisk då sammanfaller medianen och medelvärden.

**VÄNTEVÄRDET för en funktion g(X) av en s.v. X :**

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

Vi säger att en s.v. X är icke-negativ (”positiv” i boken) om $X \geq 0$.

**INTENSITETEN** för en kontinuerlig icke-negativ stokastisk variabel X definieras av

$$\lambda(x) = \frac{f(x)}{1 - F(x)}, \quad \text{för } x > 0.$$

=====================================================

ÖVNINGSUPPGIFTER

**Uppgift 1.** Den stokastiska variabeln ξ har frekvensfunktionen (täthetsfunktionen)

$$f(x) = \begin{cases} (a+1)x^2, & 0 < x < 1 \\ 0 & \textit{för övrigt} \end{cases}$$

a) Bestäm parametern $a$.
b) Beräkna P(0.1< ξ<0.3).

**Lösning:**
a)

$$\int_0^1 (a+1)x^2 dx = 1 \Rightarrow \left[\frac{(a+1)x^3}{3}\right]_0^1 = 1 \Rightarrow \frac{(a+1)}{3} = 1 \Rightarrow a = 2.$$

Därmed

$$f(x) = \begin{cases} 3x^2, & 0 < x < 1 \\ 0 & \textit{för övrigt} \end{cases}$$

b) $\int_{0.1}^{0.3} 3x^2 dx = [x^3]_{0.1}^{0.3} = 0.3^3 - 0.1^3 = 0.026.$

**Uppgift 2** En stokastisk variabel $\xi$ har täthetsfunktionen (frekvensfunktionen)

$$f(x)=\begin{cases} 0 & , \quad x<0 \\ a(10-x) & , \quad 0\le x\le 10 \\ 0 & , \quad x>10 \end{cases}$$

a) Bestäm konstanten $a$.
b) Vad är sannolikheten att $\xi > 8$ ?
c) Bestäm väntevärdet $E(\xi)$.

**Lösning:**

**a)** $\int_0^{10} a(10-x)dx = 1 \Rightarrow \left[a(10x-\frac{x^2}{2})\right]_0^{10} = 1 \Rightarrow a = \frac{1}{50}$

**b)** $P(\xi > 8) = \int_8^{10} \frac{1}{50}(10-x)dx = \left[\frac{1}{50}(10x-\frac{x^2}{2})\right]_8^{10} = \frac{1}{25} = 0.04$

**c)** $E(\xi) = \int_0^{10} x\cdot\frac{1}{50}(10-x)dx = \int_0^{10}\frac{1}{50}(10x-x^2)dx = \frac{1}{50}(5x^2-\frac{x^3}{3})\Big|_0^{10} = \frac{10}{3}$

**Svar:** **a)** $a=\frac{1}{50}$ **b)** 0.04 **c)** $\frac{10}{3}$

**Uppgift 3.**
En stokastisk variabel ξ har följande täthetsfunktion (frekvensfunktion)

$$f(x)=\begin{cases} \sin x & 0\le x\le \frac{\pi}{2} \\ 0 & \text{för övrigt.} \end{cases}$$

Bestäm väntevärdet $E(\xi)$, variansen Var(ξ) och standardavvikelsen $\sigma$.

**Lösning:**

$$\mu = E(\xi) = \int_0^{\pi/2} x\sin x dx = (\text{part. int.}) = \left[\sin x - x\cos x\right]_0^{\pi/2} = 1 .$$

För variansen använder vi formeln

$$Var(\xi) = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2 .$$

$$\int_0^{\pi/2} x^2\sin x dx = (\text{part. int. 2 gånger}) = \left[-x^2\cos x + 2x\sin x + 2\cos x\right]_0^{\pi/2} = \pi - 2 .$$

$$Var(\xi) = \int_0^{\pi/2} x^2\sin x dx - \mu^2 = \pi - 2 - 1 = \pi - 3 \approx 0.14159 .$$

Standardavvikelsen för $\xi$:

$\sigma = \sqrt{Var(\xi)}$ =0.376

**Uppgift 4.** En stokastisk variabel $\xi$ har följande fördelningsfunktion

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

Bestäm
a) medianen,
b) täthetsfunktionen (frekvensfunktionen) $f(x)$
c) väntevärdet $E(\xi)$.
d) sannolikheten $P(\ 2 \leq \xi \leq 5\ )$

**Lösning:**
**a)** Medianen är lösningen till ekvationen

$$F(x) = 0.50.$$

$1 - e^{-3x} = 1/2 \Rightarrow e^{-3x} = 1/2 \Rightarrow\ -3x = \ln(1/2) \Rightarrow\ x = \frac{\ln(1/2)}{-3} \Rightarrow\ x = \frac{\ln(2)}{3}$.

**Svar:** a) Medianen= $\frac{\ln(2)}{3}$

**b)** Frekvensfunktion $f(x) = \begin{cases} 3e^{-3x}, & x > 0 \\ 0 & x < 0 \end{cases}$ .

Anmärkning: Det är oviktigt hur vi definierar $f(0)$ eftersom $\xi$ är en kontinuerlig s.v.

**c)** $E(\xi) = \int_{-\infty}^{\infty} xf(x)dx = \int_{0}^{\infty} 3xe^{-3x}dx$

{ partiell integration
$u = 3x\ ,\quad v' = e^{-3x}$
$u' = 3\ ,\quad v = \frac{e^{-3x}}{-3}$ }

$$\int 3xe^{-3x}dx = uv - \int u'v dx = -xe^{-3x} + \int e^{-3x}dx = -xe^{-3x} - \frac{e^{-3x}}{3}.$$

Därför

$$\int_0^{\infty} 3xe^{-3x}dx = \left[-xe^{-3x} - \frac{e^{-3x}}{3}\right]_0^{\infty} = \frac{1}{3}.$$

**Svar: c)** $E(\xi) = \frac{1}{3}$

**d)** Sannolikheten $P(\ 2 \leq \xi \leq 5\ ) = F(5) - F(2)$
$= (1 - e^{-15}) - (1 - e^{-6})\ =\ e^{-6}\ - e^{-15}\ \approx 0.002478.$
**Svar: d)** 0.002478

**Uppgift 5.** Bestäm konstanten c så att

$$f(x)=\begin{cases}\dfrac{c}{\sqrt{x-2}}, & 2<x\le 5\\ 0 & \text{för övrigt}\end{cases}$$

blir en täthetsfunktion.

**Lösning:**

$f(x)$ måste satisfiera villkoret: Arean=1 dvs $\int_{-\infty}^{\infty} f(x)dx=1$.

Först beräknar vi integralen

Arean= $\int_{-\infty}^{\infty} f(x)dx=\int_{2}^{5}\frac{c}{\sqrt{x-2}}dx$

| |
|---|
| Substitution $x-2=t$, $dx=dt$,<br>Gränser : $x=2\Rightarrow t=0$,<br>$x=5\Rightarrow t=3$ |

$$=\int_0^3 \frac{c}{\sqrt{t}}dt=\int_0^3 ct^{\frac{-1}{2}}dt=\left[\frac{ct^{\frac{1}{2}}}{\frac{1}{2}}\right]_0^3=2c\sqrt{3}.$$

Från ekvationen Arean=1 har vi $2c\sqrt{3}=1\Rightarrow c=\frac{1}{2\sqrt{3}}$.

**Svar:** $c=\frac{1}{2\sqrt{3}}\approx 0.2887$

**Uppgift 6.** Rita täthetsfunktionen $f(x)$ till s.v. X och bestäm fördelningsfunktionen $F(x)$ om

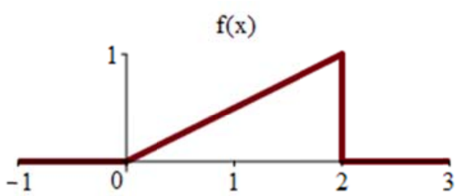
$$f(x)=\begin{cases}x/2 & \text{om } 0\le x\le 2\\ 0 & \text{för övrigt}\end{cases}$$

Rita därefter fördelningsfunktionen $F(x)$.

**Lösning:**

a) $f(x)=\begin{cases}x/2 & \text{om } 0\le x\le 2\\ 0 & \text{för övrigt}\end{cases}=\begin{cases}0 & \text{om } x<0\\ x/2 & \text{om } 0\le x\le 2\\ 0 & \text{för } x>2\end{cases}$

Grafen till täthetsfunktionen $f(x)$:

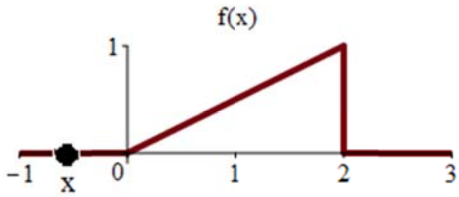


Vi bestämmer fördelningsfunktionen med hjälp av formeln $F(x)=\int_{-\infty}^{x} f(t)dt$.

Eftersom $f(x)$ är definierad med separata formler i tre intervall måste vi, vid beräkning av integralen $\int_{-\infty}^{x} f(t)dt$ betrakta tre fall:

Fall 1: $x<0$ , Fall 2: $0 \le x \le 2$ och Fall 3: $x>2$.

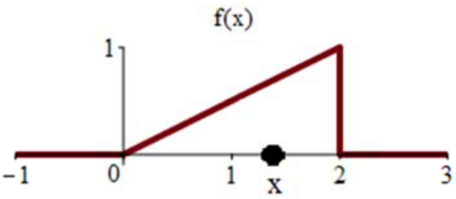
**i)** Fall1, $x<0$




Om x<0 då gäller $f(x)=0$.

Detta ger $F(x)=\int_{-\infty}^{x} f(t)dt=\int_{-\infty}^{x} 0dt=0$.

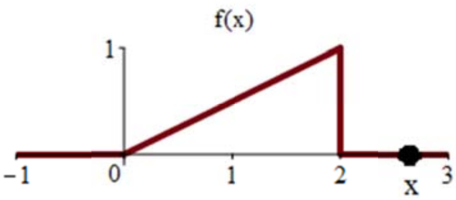
**ii)** Fall2, $0 \le x \le 2$.




Vi använder igen **samma formel** $F(x)=\int_{-\infty}^{x} f(t)dt$ men, i det här fallet, delar vi integralen $\int_{-\infty}^{x} f(t)dt$ i två delar (eftersom vi integrerar över intervallet där $f(x)$ beskrivs med två formler). Vi har

$$F(x)=\int_{-\infty}^{x} f(t)dt=\int_{-\infty}^{0} f(t)dt+\int_{0}^{x} f(t)dt=\int_{-\infty}^{0} 0dt+\int_{0}^{x} \frac{t}{2}dt=0+\left[\frac{t^2}{4}\right]_0^x=\frac{x^2}{4}.$$

iii) Fall 3), $x>2$.



I det här fallet delar vi integralen $\int_{-\infty}^{x} f(t)dt$ i tre delar (eftersom vi integrerar över intervallet där $f(x)$ beskrivs med tre formler).

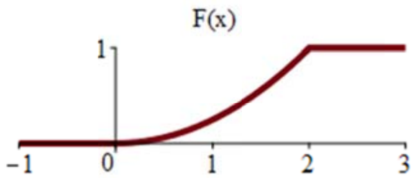
Vi har

$$F(x)=\int_{-\infty}^{x} f(t)dt=\int_{-\infty}^{0} f(t)dt+\int_{0}^{2} f(t)dt+\int_{2}^{x} f(t)dt=\int_{-\infty}^{0} 0dt+\int_{0}^{2}\frac{t}{2}dt+\int_{2}^{x} 0dt=0+\left[\frac{t^2}{4}\right]_0^2+0=1.$$

Sammanfattningsvis har vi

$$F(x)=\begin{cases} 0 & \text{om } x<0 \\ \frac{x^2}{4} & \text{om } 0\le x\le 2 \\ 1 & \text{om } x>2 \end{cases}$$

Grafen till fördelningsfunktionen $F(x)$.



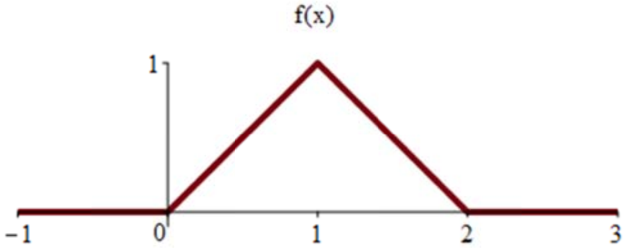

**Uppgift 7.** Rita täthetsfunktionen $f(x)$ till s.v. X och bestäm fördelningsfunktionen $F(x)$ om

$$f(x)=\begin{cases} x & \text{om } 0\le x\le 1 \\ 2-x & \text{om } 1< x\le 2 \\ 0 & \text{för } \textit{övrigt} \end{cases}$$

Rita därefter fördelningsfunktionen $F(x)$.

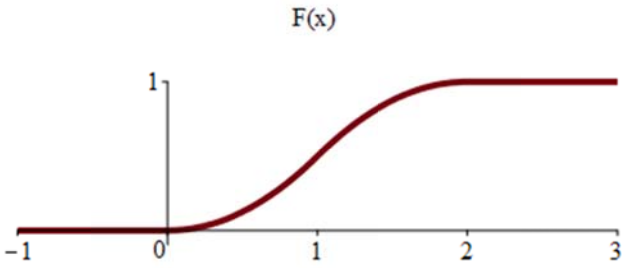
**Svar:**

i) Grafen till täthetsfunktionen $f(x)$:



ii) Fördelningsfunktionen $F(x)=\begin{cases} 0 & \text{om } x<0 \\ \dfrac{x^2}{2} & \text{om } 0\le x\le 1 \\ -1+2x-\dfrac{x^2}{2} & \text{om } 1<x\le 2 \\ 1 & \text{om } x>2 \end{cases}$

iii) Grafen till fördelningsfunktionen $F(x)$.




**Uppgift 8.** En stokastisk variabel X har fördelningsfunktionen

$$F(x)=\begin{cases} 1-\dfrac{4}{x^2} & \text{om } x\ge 2 \\ 0 & x<2. \end{cases}$$

Beräkna **a)** medianen och **b)** väntevärdet (medelvärdet) till s.v. X.

**Lösning:**

a) Medianen bestämmer vi genom att lösa en av följande ekvationer:

$F(x)=0.5$ eller $\int_{-\infty}^{x} f(x)dx=0.5$.

I vårt fall är det enklare att lösa

$$F(x)=0.5 \Rightarrow 1-\frac{4}{x^2}=0.5 \Rightarrow \frac{4}{x^2}=\frac{1}{2} \Rightarrow x^2=8 \Rightarrow x=\pm\sqrt{8}.$$

Eftersom $x\ge 2$ väljer vi $x=\sqrt{8}\approx 2.828$.

**b)** Medelvärdet till s.v. X. beräknas med hjälp av följande formel

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

Först måste vi bestämma täthetsfunktionen

$$f(x) = F'(x) = \begin{cases} 8x^{-3} & \text{om x} > 2 \\ 0 & x < 2. \end{cases}$$

Nu kan vi beräkna

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_{2}^{\infty} x \cdot 8x^{-3}dx = \int_{2}^{\infty} 8x^{-2}dx = \left[\frac{-8}{x}\right]_2^{\infty} = -0 + 4 = 4 .$$

**Svar**: a) medianen är $\sqrt{8} \approx 2.828$ b) väntevärdet (=medelvärdet)=4

**Uppgift 9.** Livslängden hos en viss transistortyp är exponentialfördelad s.v. med fördelningsfunktionen $F(x) = \begin{cases} 1 - e^{-x/2} & \text{om x} \geq 0 \\ 0 & x < 0. \end{cases}$

**a)** Bestäm sannolikheten att en sådan transistor (slumpvis vald) har livslängden som är större än 1 år.
**b)** Man köper 5 transistorer. Bestäm sannolikheten att minst 4 av dem har livslängden som är större än 1 år.

**Lösning:**

a) $p = P(X > 1) = 1 - P(X \leq 1) = 1 - F(x) = 1 - (1 - e^{-1/2}) = e^{-1/2} \approx 0.60653.$
b) Låt Y beteckna antalet transistorer bland dem 5 köpta som har livslängden större än 1 år .
Då är $Y \in Bin(5, p)$ där $p \approx 0.60653$ och $q = 1 - p \approx 0.39347$ .

$$P(Y \geq 4) = P(Y = 4) + P(Y = 5) = \binom{5}{4} p^4 q + \binom{5}{5} p^5 q^0 = 5 \cdot p^4 q + p^5$$

$$= 0.26625 + 0.08208 = 0.3483 .$$

**Svar:** a) 0.60653
b) $\binom{5}{4} p^4 q + \binom{5}{5} p^5 q^0 = 5 \cdot p^4 q + p^5 = 0.3483$

**Uppgift 10.** Den s.v. X har täthetsfunktionen $f(x) = 3e^{-3x}, x \geq 0$.
Beräkna väntevärdet $E(g(X))$ där $g(x) = e^{-2x}$.

**Lösning:**

$$E(g(X)) = \int_{-\infty}^{\infty} g(x) f(x)dx = \int_{0}^{\infty} e^{-2x} \cdot 3e^{-3x}dx = \int_{0}^{\infty} 3e^{-5x}dx = \left[3\frac{e^{-5x}}{-5}\right]_0^{\infty} = 0 - (-\frac{3}{5}) = \frac{3}{5}.$$

**Svar:** 3/5