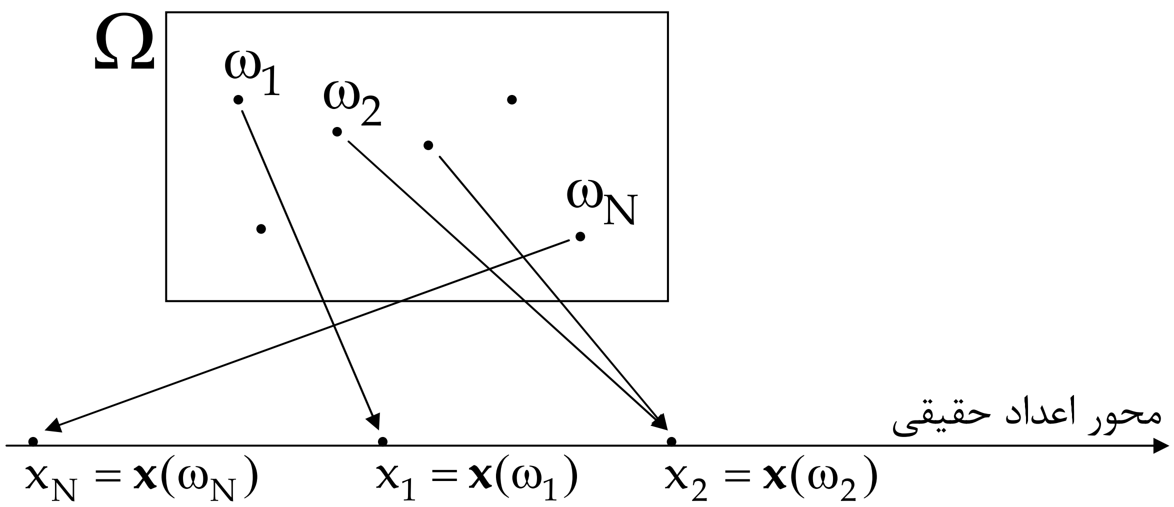
فصل ۴: متغیر تصادفی Random Variable (RV)

Chapter 4
Chapter 5: Section 5.2

۱. تعریف متغیر تصادفی
۲. تابع pmf
۳. تابع CDF
۴. تابع pdf
۵. برخی توزیع‌های خاص
۶. تابع یک متغیر تصادفی
۷. میانگین و واریانس یک متغیر تصادفی
۸. میانگین و واریانس برخی توزیع‌های خاص
۹. گشتاورهای یک متغیر تصادفی

متغیر تصادفی تابعی است از نقاط فضای نمونه مثل $\mathbf{x}(\omega)$ که به هر یک از نقاط فضای نمونه عددی را نسبت می‌دهد.

دامنهٔ (Domain) این تابع مجموعهٔ Ω است و برد (Range) آن مجموعه‌ای از اعداد حقیقی است.



(متغیر تصادفی حقیقی که مورد بحث ما است، عددی حقیقی را نسبت می‌دهد. متغیر تصادفی مختلط، عددی مختلط را نسبت می‌دهد.)

چون خیلی اوقات به جای نتایج آزمایش، تابعی از نتایج مورد توجه ما است. مثلاً در پرتاب دو تاس مقدار حاصلجمع اعداد دو تاس یا در پرتاب سکه‌ها، مجموع تعداد شیرهای ظاهر شده ممکن است مد نظر باشد.

مثال: در آزمایش پرتاب سه سکهٔ سالم داریم:

$$\Omega = \{\underbrace{HHH}_{\omega_1},\ldots,\underbrace{TTT}_{\omega_8}\}$$

مثلاً تعداد شیرها یک متغیر تصادفی است:

$\mathbf{x}(\omega) = \omega$ تعداد شیرها در واقعهٔ

ω	HHH	HHT	HTH	HTT	THH	THT	TTH	TTT
$\mathbf{x}(\omega)$	۳	۲	۲	۱	۲	۱	۱	۰

حال در مثال فوق احتمال اینکه $\mathbf{x}(\omega) \le 1$ باشد چیست؟ احتمال روی واقعه‌ها تعریف می‌شود. احتمال $\mathbf{x}(\omega) \le 1$ یعنی احتمال مجموعهٔ آن ωهایی که برای آنها $\mathbf{x}(\omega) \le 1$ باشد:

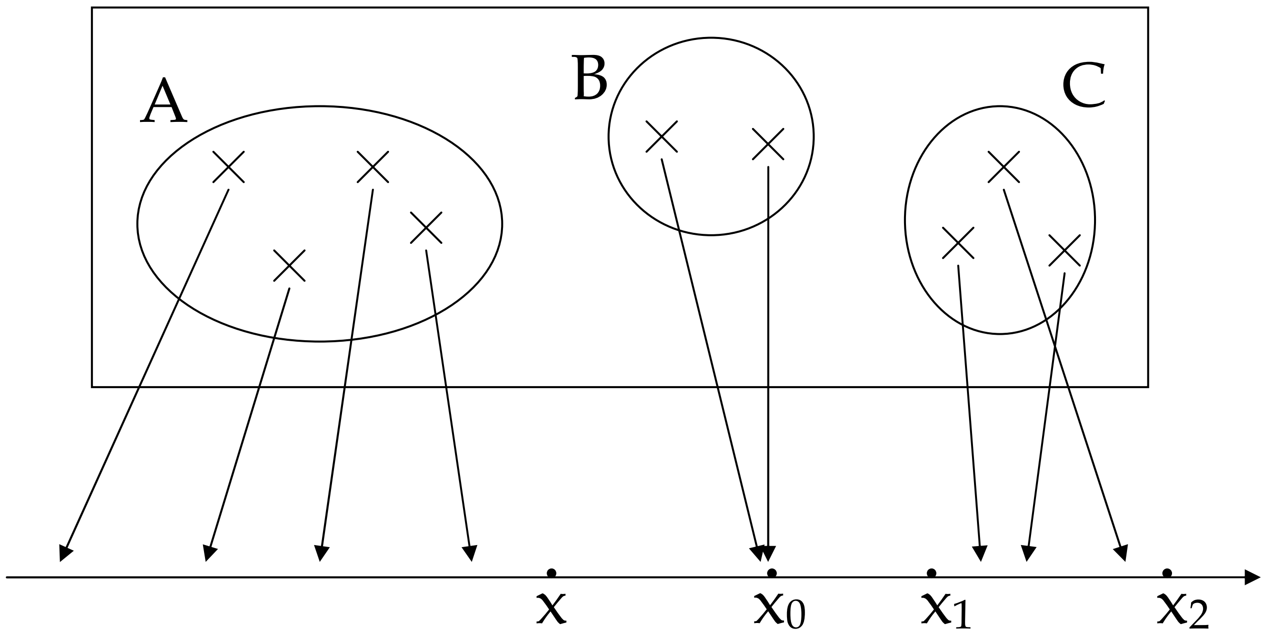
$$P\{\mathbf{x} \le 1\} = P\{\omega : \mathbf{x}(\omega) \le 1\} = P\{HTT, THT, TTH, TTT\} = \frac{4}{8}$$

به طور خلاصه می‌نویسیم: $\{\mathbf{x} \le x\}$، ولی این مجموعه‌ای از اعداد نیست بلکه یعنی: $\{\omega : \mathbf{x}(\omega) \le x\}$.

احتمال اینکه $\mathbf{x}(\omega) = 2$ باشد چیست؟

$$P\{\mathbf{x} = 2\} = P\{\omega : \mathbf{x}(\omega) = 2\} = P\{HHT, HTH, THH\} = \frac{3}{8} = \binom{3}{2}(\frac{1}{2})^2(\frac{1}{2})^{3-2}$$

یعنی:



$$A = \{\mathbf{x} \le x\}$$
$$B = \{\mathbf{x} = x_0\}$$
$$C = \{x_1 < \mathbf{x} \le x_2\}$$

حال مثلاً اگر احتمال $\{\mathbf{x} \le x\}$ را برای هر x بدانیم، احتمال همهٔ واقعه‌های مورد نظر را خواهیم داشت و لزومی به دانستن فضای احتمال Ω نیست.

$-\infty$ یا $+\infty$ اعداد حقیقی نیستند، لذا $\mathbf{x}(\omega)$ نباید $-\infty$ یا $+\infty$ شود (مگر آنکه احتمال آن ω صفر باشد). یعنی اجازه می‌دهیم که $\mathbf{x}(\omega)$ برای برخی ωها $-\infty$ یا $+\infty$ شود مشروط بر آنکه احتمال آن ωها صفر باشد:

$$P\{\mathbf{x} = +\infty\} = P\{\mathbf{x} = -\infty\} = 0$$

(تعریف: متغیر تصادفی حقیقی $\mathbf{x}(\omega)$ تابعی است حقیقی از نقاط فضای نمونهٔ $\omega \in \Omega$ به طوری که برای هر عدد حقیقی x، مجموعهٔ $\{\mathbf{x}(\omega) \le x\}$ یک واقعه باشد و $P\{\mathbf{x} = +\infty\} = P\{\mathbf{x} = -\infty\} = 0$)

متغیر تصادفی را گسسته گویند هرگاه مقادیری که $\mathbf{x}(\omega)$ می‌تواند اختیار کند (برد تابع) قابل شمارش باشد.

در حالت گسسته مثل مثالی که داشتیم ساده‌ترین راه برای مشخص کردن احتمال واقعه‌ها، مشخص کردن احتمال واقعه‌های $\{\mathbf{x} = x_i\}$ برای x_iهای ممکنه است (نسبت به مشخص کردن احتمال $\{\mathbf{x} \le x\}$ برای هر x).

تابع جرمی احتمال pmf (Probability Mass Function) یا تابع احتمال یا تابع فراوانی:

تعریف: اگر $\mathbf{x}$ فقط مقادیر ممکن x_1، x_2، x_3 و ... (قابل شمارش) را با احتمالهای p_1، p_2، p_3 و ... اختیار کند، تابع احتمال متغیر تصادفی $\mathbf{x}$ به صورت زیر تعریف می‌شود:

آرگومان تابع

$$\underset{\substack{\downarrow \\ \text{بیانگر نوع تابعیت}}}{P_{\mathbf{x}}}\overset{\substack{\text{آرگومان تابع} \\ \uparrow}}{(x)} = \text{Prob}\{\mathbf{x} = x\} = \begin{cases} p_i & x = x_i \\ 0 & \text{otherwise} \end{cases}$$

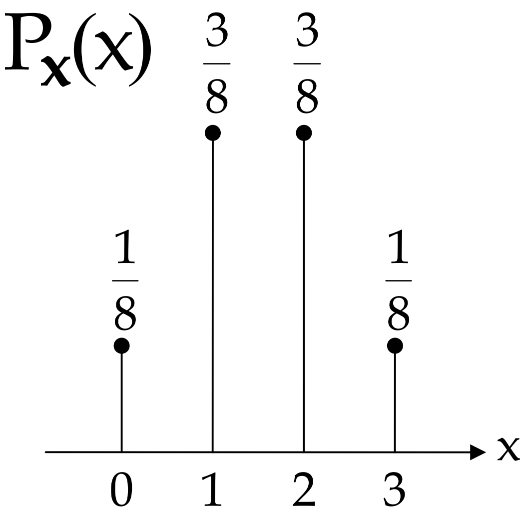
بیانگر نوع تابعیت

(اگر فقط با یک متغیر تصادفی سروکار داشته باشیم اندیس $\mathbf{x}$ را می‌توان حذف کرد: $P(x_i)$)

در مثالی که داشتیم، برای i = 0, 1, 2, 3 داریم:

$$P_{\mathbf{x}}(i) = \text{Prob}\{\mathbf{x} = i\} = \binom{3}{i}\left(\frac{1}{2}\right)^{i}\left(\frac{1}{2}\right)^{3-i} = \frac{\binom{3}{i}}{8} = \begin{cases} \frac{1}{8} & i = 0 \\ \frac{3}{8} & i = 1 \\ \frac{3}{8} & i = 2 \\ \frac{1}{8} & i = 3 \end{cases}$$

و برای سایر xها داریم: $P_{\mathbf{x}}(x) = 0$.



تابع $P_{\mathbf{x}}(x)$ است که فقط در یک سری نقاط محدود یا حداکثر قابل شمارش مقدار غیرصفر دارد.

به طور کلی روشن است که $P(x_i)$ها اعدادی بین صفر و یک هستند (چون احتمال واقعه‌های $\{\mathbf{x}=x_i\}$اند) و داریم:

$$\sum_{i=1}^{\infty} P(x_i)=1$$

اگر احتمال هر واقعه را بخواهیم، با داشتن $P(x_i)$ها قابل محاسبه است. مثلاً در مثال فوق داریم:

$$P\{\mathbf{x}\le 1.1\}=P\{\mathbf{x}=0\}+P\{\mathbf{x}=1\}=\frac{1}{8}+\frac{3}{8}=\frac{4}{8}$$

و به طور کلی برای هر زیرمجموعه از اعداد حقیقی مثل A داریم:

$$\mathrm{Prob}\{\mathbf{x}\in A\}=\sum_{x_i\in A} P(x_i)$$

تابع احتمال برای متغیر تصادفی گسسته تعریف می‌شود.

(وقتی متغیر تصادفی پیوسته باشد یعنی مقادیری که می‌تواند اختیار کند غیرقابل شمارش باشد، تعیین احتمال $\{\mathbf{x}=x_i\}$ها کفایت نمی‌کند. اغلب همه صفرند. در اینجا می‌توانیم تابع توزیع انباشته را مطرح کنیم.)

در مهندسی برق خیلی اوقات با متغیرهای تصادفی پیوسته مثل ولتاژ، توان و امثالهم سروکار داریم.

تابع توزیع انباشته (CDF (Cumulative Distribution Function:

طبق تعریف داریم:

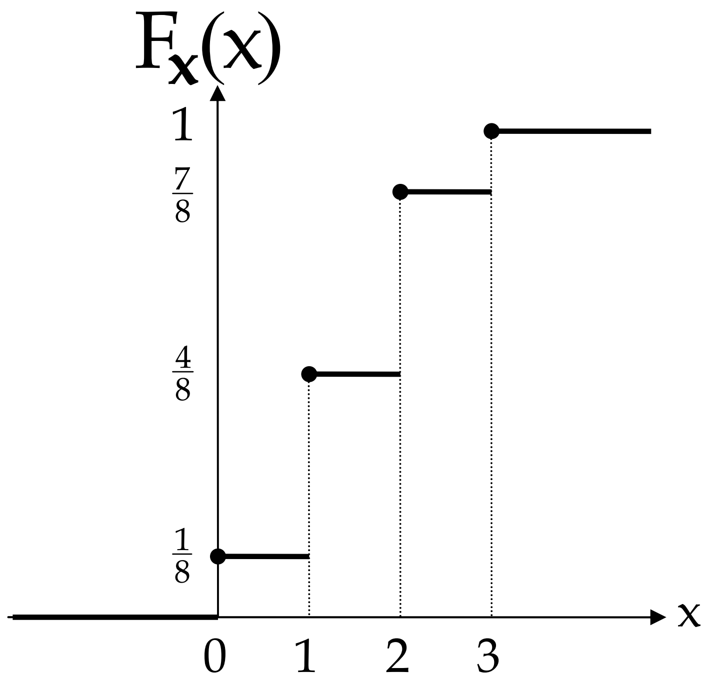
آرگومان تابع

$$\underset{\text{بیانگر نوع تابعیت}}{F_{\mathbf{x}}}(\overset{\uparrow}{x}) = \text{Prob}\{\mathbf{x} \le x\}$$

(اگر فقط با توزیع یک متغیر تصادفی سروکار داشته باشیم، $F(x)$ کفایت می‌کند.)

چه برای متغیر تصادفی پیوسته و چه گسسته اگر احتمال واقعه‌های $\{\mathbf{x} \le x\}$ را برای هر x بدانیم، احتمال همهٔ واقعه‌ها مشخص خواهد شد. پس $F(x)$ به طور کامل می‌تواند متغیر تصادفی را توصیف کند.

مثال ۱: در مثالی که داشتیم:



$$F_{\mathbf{x}}(-0.001) = P\{\mathbf{x} \le -0.001\} = 0$$
$$F_{\mathbf{x}}(0) = P\{\mathbf{x} \le 0\} = \tfrac{1}{8}$$
$$F_{\mathbf{x}}(0.001) = P\{\mathbf{x} \le 0.001\} = \tfrac{1}{8}$$
$$F_{\mathbf{x}}(0.999) = P\{\mathbf{x} \le 0.999\} = \tfrac{1}{8}$$
$$F_{\mathbf{x}}(1) = P\{\mathbf{x} \le 1\} = \tfrac{4}{8}$$
$$F_{\mathbf{x}}(1.0001) = P\{\mathbf{x} \le 1.0001\} = \tfrac{4}{8}$$
$$F_{\mathbf{x}}(10) = P\{\mathbf{x} \le 10\} = 1$$

در حالت گسسته $F_{\mathbf{x}}(x)$ به صورت پلکانی است.

می‌دانیم که در مورد RV گسسته: $P\{\mathbf{x} \in A\} = \sum_{x_i \in A} P(x_i)$؛ لذا داریم:

$$F(x) = P\{\mathbf{x} \le x\} = \sum_{x_i \le x} P(x_i) = \sum_{i} P(x_i) u(x - x_i)$$

که:

$$u(x) = \begin{cases} 1 & x \ge 0 \\ 0 & x < 0 \end{cases}$$

و همچنین:

$$P(x_k) = x_k \text{ در } F(x) \text{ مقدار پرش } = F(x_k) - F(x_k^-)$$

(به طور ریاضی نیز این را نشان خواهیم داد.)

مثال ۲: یک مکالمهٔ تلفنی در زمانی کاملاً تصادفی بین $[0,T]$ واقع می‌شود.

$$\Omega=\{t:0\le t\le T\}$$

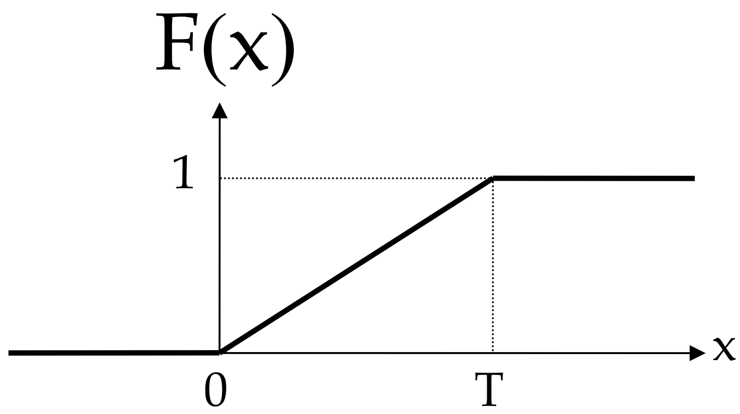
به طوری که برای هر $0\le t_1\le t_2\le T$:

$$P\{t_1\le t\le t_2\}=\frac{t_2-t_1}{T}$$

اگر متغیر تصادفی ما، خود زمان وقوع مکالمهٔ تلفنی باشد، داریم:

$$\mathbf{x}(t)=t:0\le t\le T$$

$$F_{\mathbf{x}}(x)=P\{\mathbf{x}(t)\le x\}=\begin{cases}P(\Omega)=1 & x>T\\ P\{0\le t\le x\}=\dfrac{x}{T} & 0\le x\le T\\ P(\varnothing)=0 & x<0\end{cases}$$



تابع توزیع انباشتهٔ Ramp برای توزیع یکنواخت حاصل می‌شود.

توجه کنید که عناصر Ω اعداد بین 0 تا T بودند، ولی $F(x)$ برای همهٔ xها تعریف شده است (مثلاً xهای منفی متناظر با مجموعهٔ $\varnothing$ می‌شوند و احتمال صفر دارند).

خواص تابع توزیع انباشته (CDF):

$$1)\ F(-\infty)=0$$

زیرا:

$$F(-\infty)=P\{\mathbf{x}=-\infty\}=0$$

$$2)\ F(+\infty)=1$$

زیرا:

$$F(+\infty)=P\{\mathbf{x}\le+\infty\}=1$$

3) $x_1<x_2\Rightarrow F(x_1)\le F(x_2)$: تابعی غیرنزولی است $F(x)$

زیرا:

$$x_1<x_2\Rightarrow\{\omega:\mathbf{x}(\omega)\le x_1\}\subset\{\omega:\mathbf{x}(\omega)\le x_2\}\Rightarrow P\{\mathbf{x}\le x_1\}\le P\{\mathbf{x}\le x_2\}\Rightarrow F(x_1)\le F(x_2)$$

از اینجا نتیجه می‌شود که: $0\le F(x)\le 1$، زیرا دیدیم که در $-\infty$، صفر و در $+\infty$، یک است.

4) $P\{\mathbf{x} > x\} = 1 - F(x)$

زیرا:

$$\{\omega : \mathbf{x}(\omega) > x\} = \{\omega : \mathbf{x}(\omega) \le x\}^c \Rightarrow P\{\mathbf{x} > x\} = 1 - P\{\mathbf{x} \le x\} = 1 - F(x) = F^c(x)$$

اغلب در مخابرات با $P\{\mathbf{x} > x\}$ سروکار داریم (بزرگتر شدن از یک سطح آستانه) که مکمل $F(x)$ می‌شود.

5) $P\{x_1 < \mathbf{x} \le x_2\} = F(x_2) - F(x_1)$

زیرا:

$$\{\mathbf{x} \le x_2\} = \{\mathbf{x} \le x_1\} \cup \{x_1 < \mathbf{x} \le x_2\}$$

دو مجموعهٔ جدا از هم

$$\Rightarrow \underbrace{P\{\mathbf{x} \le x_2\}}_{F(x_2)} = \underbrace{P\{\mathbf{x} \le x_1\}}_{F(x_1)} + P\{x_1 < \mathbf{x} \le x_2\}$$

$$\Rightarrow P\{x_1 < \mathbf{x} \le x_2\} = F(x_2) - F(x_1)$$

از همین جا به ازای $x_2 = x$ و $x_1 = x - \varepsilon$ و میل دادن ε به سمت صفر نتیجه می‌شود که:

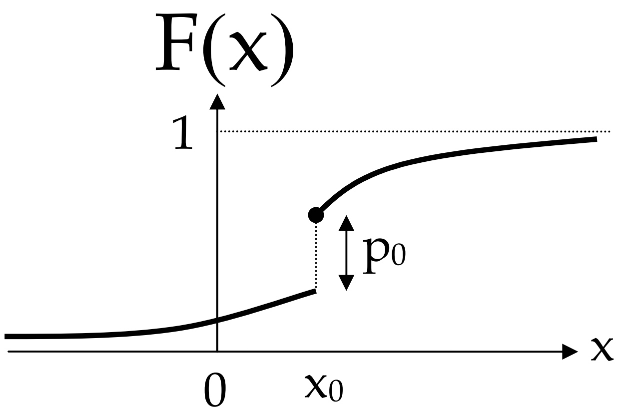
$$P\{\mathbf{x} = x\} = F(x) - F(x^-)$$

اگر پرش در $F(x)$ داشته باشیم، $F(x)$ با $F(x^-)$ به اندازهٔ $P(x)=P\{\mathbf{x}=x\}$ متفاوت خواهد بود، ولی $F(x)$ به هر حال از راست پیوسته است، یعنی:

$$F(x_0)=F(x_0^+)=P\{\mathbf{x}\le x_0\}$$

$$F(x_0^-)=P\{\mathbf{x}<x_0\}$$

$$p_0=F(x_0^+)-F(x_0^-)=F(x_0)-F(x_0^-)$$



(ضمناً با توجه به $P\{\mathbf{x}=x\}=F(x)-F(x^-)$ داریم:

$P\{x_1\le\mathbf{x}\le x_2\}=F(x_2)-F(x_1^-)$)

6) $F(x)=F(x^+)$

(قابل اثبات به طور دقیق ریاضی، کتاب DeGroot، ص ۱۱۰)

حال که $F(x)$ را فهمیدیم می‌توانیم گسسته یا پیوسته بودن متغیر تصادفی را طور دیگری نیز تعریف کنیم:

متغیر تصادفی را گسسته (Discrete) گویند اگر $F(x)$ به صورت پلکانی باشد، مانند مثالی که برای سه سکه داشتیم.

در این حالت همان طور که قبلاً دیدیم:

$$P\{\mathbf{x}=x_i\}=F(x_i)-F(x_i^-)=p_i=P(x_i)$$

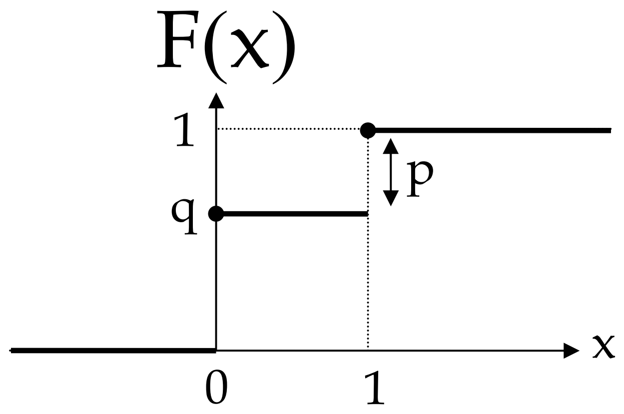
تعریف فوق با این تعریف که مقادیر ممکنه برای $\mathbf{x}$ قابل شمارش هستند معادل است.

توجه کنید که گر چه اگر فضای نمونه (Ω) قابل شمارش باشد، هر $\mathbf{x}$ای روی آن گسسته خواهد بود، ولی روی فضای Ω پیوسته نیز می‌توان $\mathbf{x}$ گسسته تعریف کرد.

مثلاً اگر A واقعه‌ای در فضای نمونهٔ غیرقابل شمارش Ω باشد و تعریف کنیم:

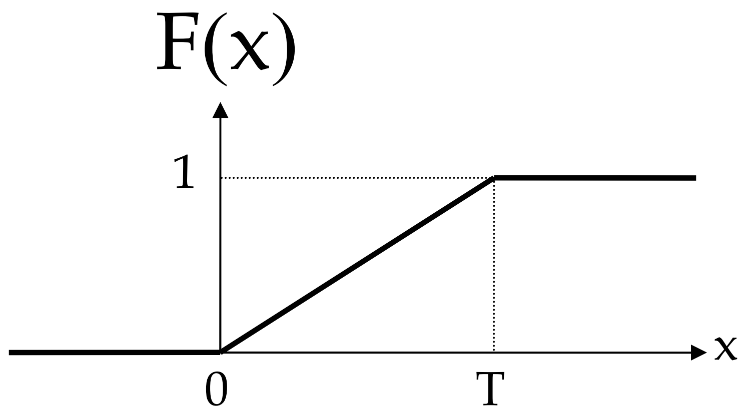
متغیر تصادفی یک-صفر (Zero-One RV) متناظر با واقعهٔ A: $\mathbf{x}(\omega)=\begin{cases}1 & \omega\in A\\ 0 & \omega\notin A\end{cases}$

در این صورت اگر $P(A)=p$ و $P(\bar{A})=1-p=q$ باشد، داریم:



متغیر تصادفی را پیوسته (Continuous) گویند اگر $F(x)$ برای هر x پیوسته باشد.

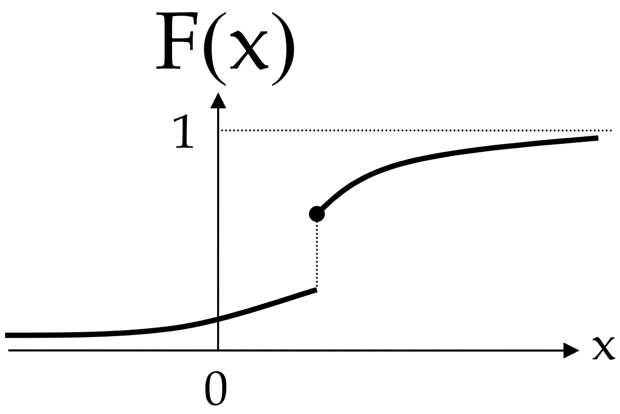
مانند مثالی که داشتیم (تلفن):



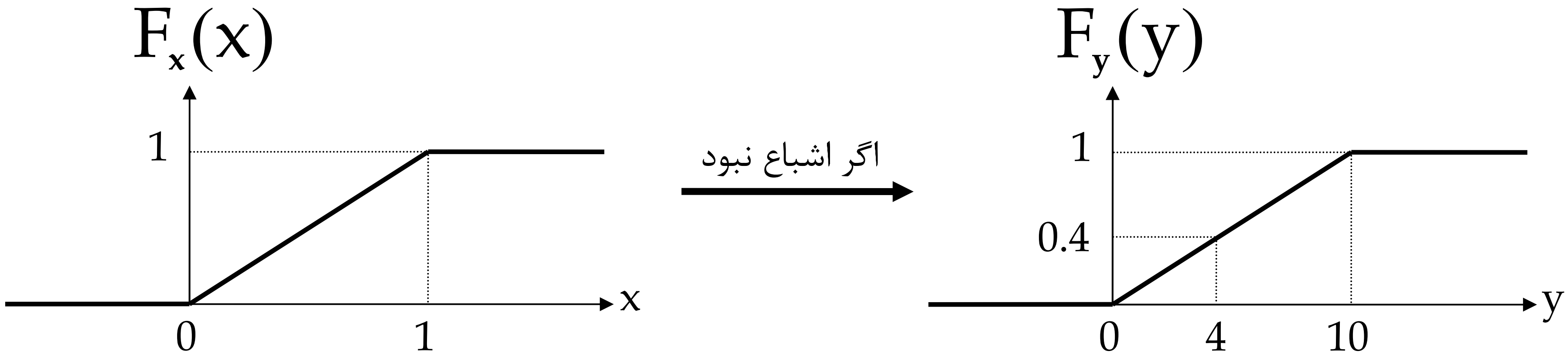
با توجه به ویژگی ۵ برای متغیر تصادفی پیوسته، نتیجه می‌گیریم که برای هر x داریم: $P\{\mathbf{x} = x\} = 0$.

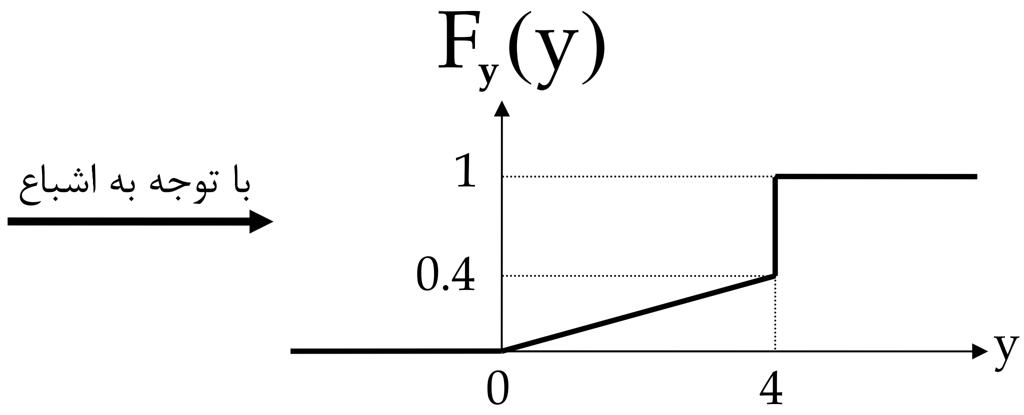
متغیر تصادفی را از نوع مخلوط (Mixed) گویند اگر $F(x)$ دارای ناپیوستگی باشد، ولی پلکانی نباشد.

(تعریفی که قبلاً برای متغیر تصادفی پیوسته کردیم دو نوع اخیر را با هم در برمی‌گرفت، چون در هر دو تعداد مقادیر ممکنه برای $\mathbf{x}$ غیرقابل‌شمارش است.)



مثال: ولتاژ ورودی تقویت‌کننده‌ای، ولتاژی کاملاً تصادفی بین 0 و 1 (احتمال یکنواخت) است. بهرهٔ تقویت‌کننده ۱۰ است، ولی حداکثر ولتاژ خروجی V ۴ است و برای بیشتر از آن روی V ۴ اشباع می‌شود.





$\{\mathbf{x}=4\}$ احتمالی برابر با ۰/۶ دارد و سایر $\{\mathbf{x}=x\}$ها احتمالشان صفر است.

صدک‌ها (نقاط درصد) (Percentiles):

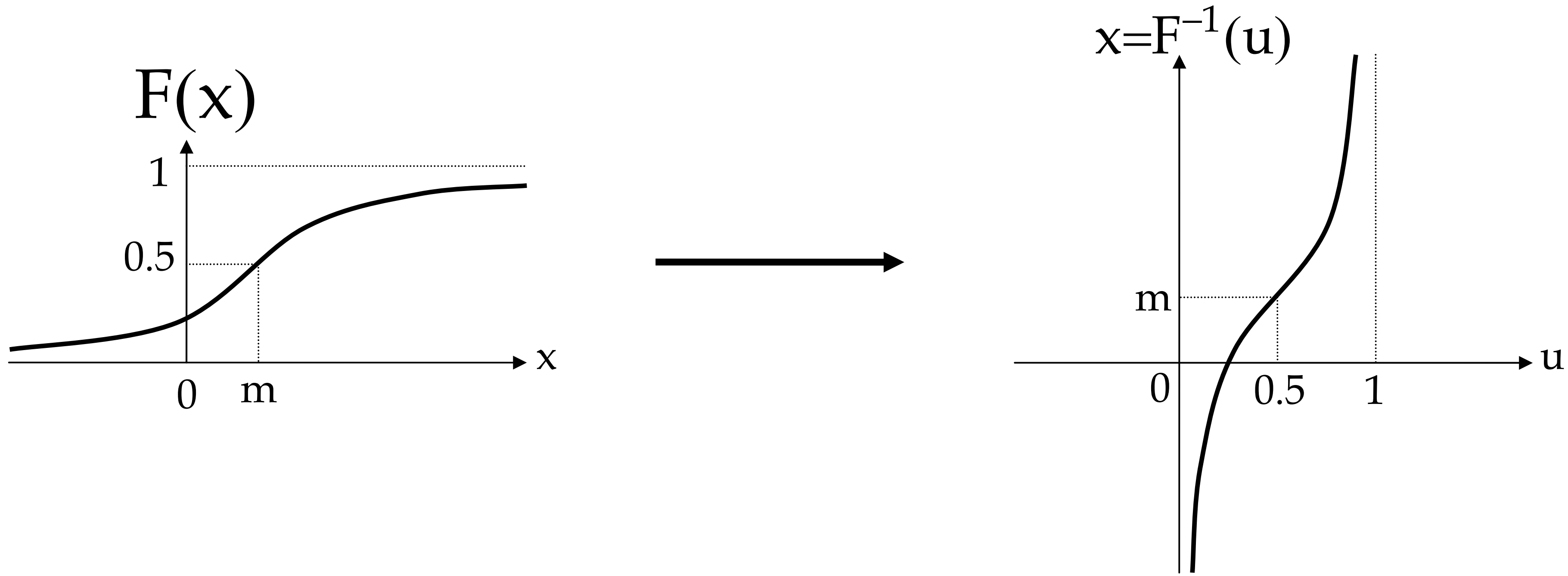
وقتی می‌گوییم $F_{\mathbf{x}}(x)=u$ یعنی به احتمال u متغیر تصادفی $\mathbf{x}$ از مقدار x بزرگتر نمی‌شود (کوچکتر یا مساوی x است). خیلی اوقات با عکس مسأله مواجهیم. مثلاً می‌خواهیم ببینیم متغیر تصادفی $\mathbf{x}$ به احتمال ۰/۹۵ کوچکتر یا مساوی کدام x است؟

$$F(x)=P\{\mathbf{x}\le x\}=u$$

$0\le u\le 1$ داده شده و $-\infty\le x\le +\infty$ مطلوب است:

تابعی از u: $x_u=F^{-1}(u)$

x_u را صدک u-۱۰۰ام متغیر تصادفی $\mathbf{x}$ گویند. مثلاً $x_{0.95}$، صدک نود و پنجم است.



صدک دهم را دهک (Decile) اول، صدک بیستم را دهک دوم و ... گویند.

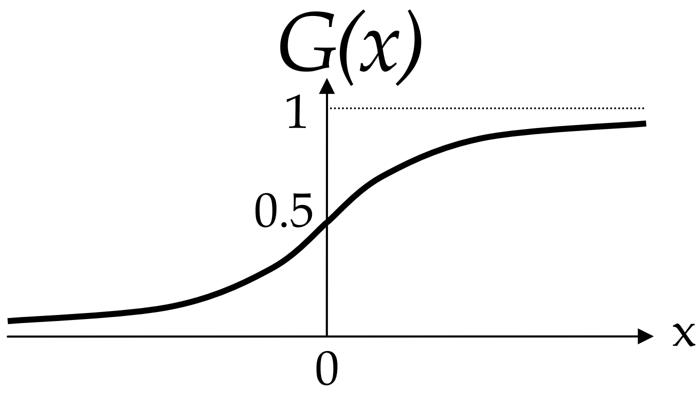
صدک ۲۵ام را چارک (Quartile) اول یا پایین و صدک ۷۵ام را چارک سوم یا بالا گویند.

مهمتر از همه اینکه صدک پنجاهم را چارک دوم یا میانه (Median) گویند.

پس میانهٔ (m) متغیر تصادفی **x** مقداری است که به ازای آن داریم:

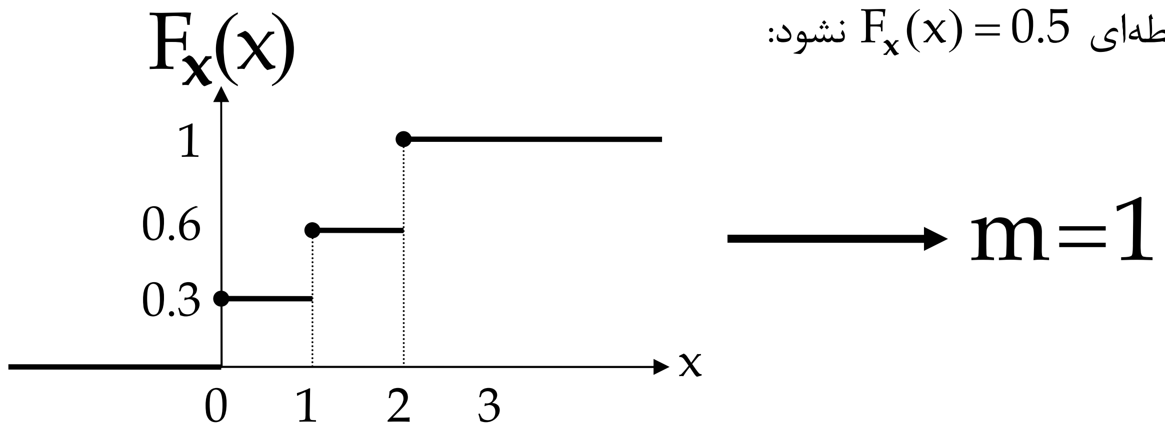
$$F_{\mathbf{x}}(m) = 0.5$$

مثال: $F_{\mathbf{x}}(x) = G(x)$؛

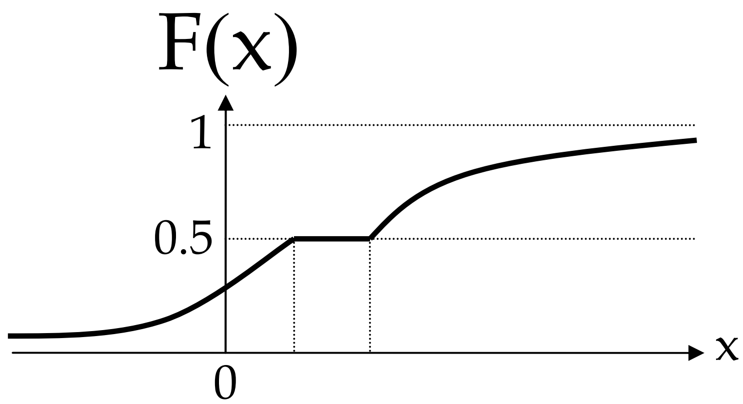


$$\longrightarrow \quad m = 0$$

ولی ممکن است F پرش داشته باشد و لذا در هیچ نقطه‌ای $F_x(x) = 0.5$ نشود:



یا ممکن است یک بازه میانه باشد:

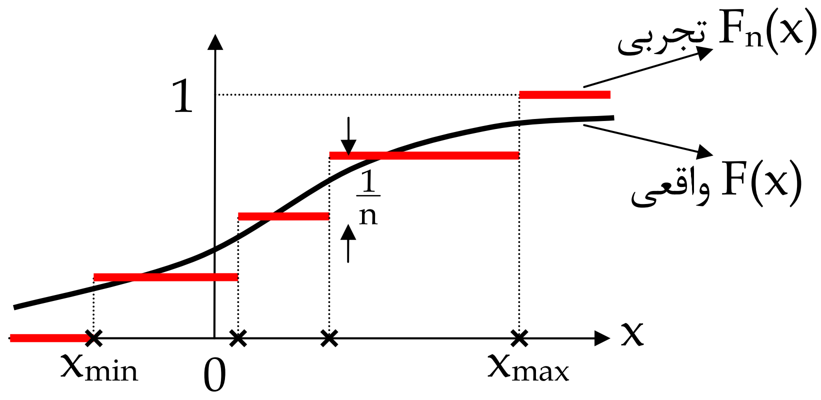


تعریف میانه: برای متغیر تصادفی $\mathbf{x}$، میانهٔ m مقداری است که برای آن داشته باشیم:

$$P\{\mathbf{x} \geq m\} \geq \frac{1}{2}, P\{\mathbf{x} \leq m\} \geq \frac{1}{2}$$

به دست آوردن تجربی F(x):

آزمایش تصادفی را nبار تکرار می‌کنیم و مقادیر حاصله برای متغیر تصادفی $\mathbf{x}$ در این nبار $(x_i : i = 1, 2, \ldots, n)$ را مرتب می‌کنیم.



مثلاً با قرائت مکرر ولتاژ نویزی که توزیع آن را نمی‌دانیم می‌توانیم F(x) تقریبی آن را به دست آوریم:

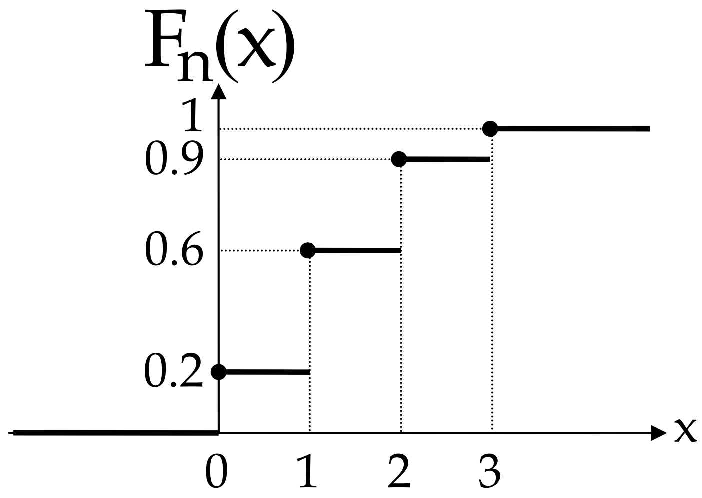
n_x تعداد x_iهایی است که برای آنها $x_i \leq x$ است: $F(x) \simeq F_n(x) = \frac{n_x}{n}$

برای n بزرگ منحنی تجربی به واقعی میل می‌کند.

اگر یک مقدار بیش از یک بار حاصل شود (در مورد متغیر تصادفی گسسته معمولاً این طور است)، پله به جای $\frac{1}{n}$، $\frac{k}{n}$ خواهد بود.

مثلاً در مثال سه سکه که داشتیم اگر در ۱۰ بار تکرار آزمایش، این نتایج حاصل شده باشند، داریم:

تعداد شیرها	صفر	۱	۲	۳
تعداد وقوع	دو بار	چهار بار	سه بار	یک بار



مقادیر واقعی:

$$F(0)=\tfrac{1}{8} \quad F(1)=\tfrac{4}{8} \quad F(2)=\tfrac{7}{8} \quad F(3)=1$$

(به صورت مشابهی منحنی صدک نیز به طور تجربی قابل تخمین است.)

تابع چگالی احتمال (pdf (Probability Density Function:

برای متغیر تصادفی پیوسته راه دیگر برای مشخص کردن احتمال واقعه‌هایی که متغیر تصادفی **x** را تعریف می‌کنند آن است که (مشابه آنچه مستقیماً برای احتمال خود واقعه‌های Ω با تعریف $\alpha(x)$ کردیم) تابعی که بیانگر چگالی احتمال است را داشته باشیم.

(به وسیلهٔ pdf دید بیشتری (در مقایسه با CDF) نسبت به میزان محتمل بودن مقادیر مختلف پیدا می‌کنیم.)

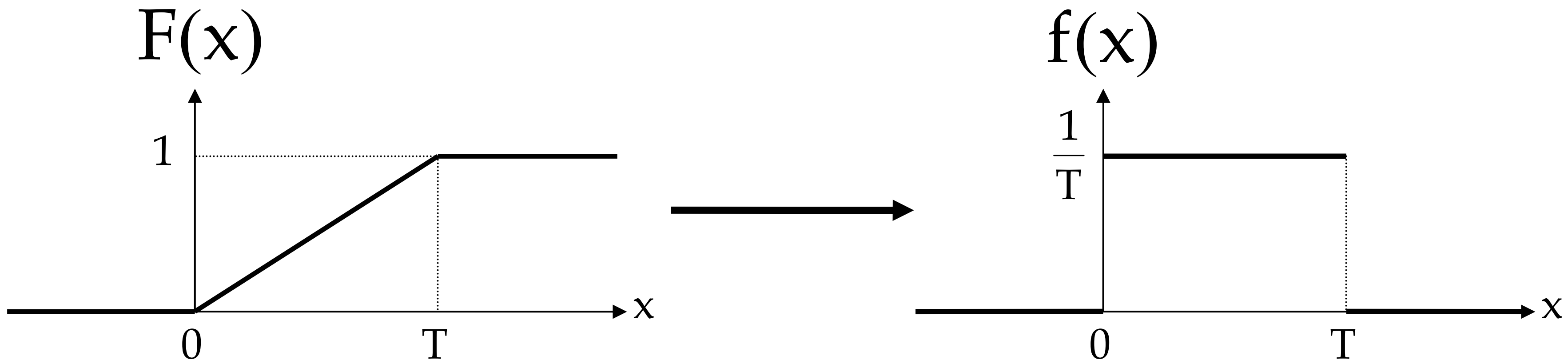
pdf را به صورت زیر تعریف می‌کنیم:

$$f_{\mathbf{x}}(x) = \frac{dF_{\mathbf{x}}(x)}{dx}$$

در کتاب نتاسیون (Notation) آن با تابع احتمال یکی است.

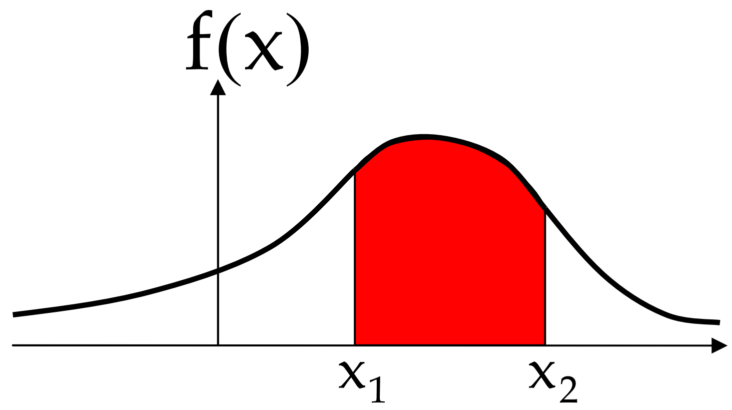
گاهی CDF و گاهی pdf را تابع توزیع (بیشتر CDF) یا توزیع (بیشتر pdf) متغیر تصادفی **x** گویند.

مثال: در مثال مکالمهٔ تلفنی کاملاً تصادفی بین $[0,T]$ دیدیم که:



خواص تابع چگالی احتمال (pdf):

1) $P\{x_1 \le \mathbf{x} \le x_2\} = \int_{x_1}^{x_2} f_{\mathbf{x}}(x)\,dx$



(یعنی واقعاً تابع چگالی احتمال است)

اثبات: می‌دانیم که:

$$P\{x_1 \le \mathbf{x} \le x_2\} = F(x_2) - F(x_1)$$

و از تعریف pdf داریم:

$$F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(x)\,dx$$

پس حکم ثابت است (توجه کنید که متغیر تصادفی را پیوسته فرض کردیم).

ضمناً اگر در رابطهٔ فوق $x_1 = -\infty$ قرار دهیم، داریم:

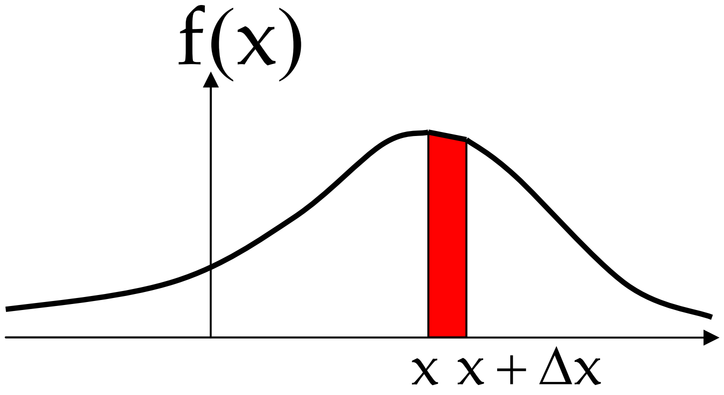
$$F(x) = \int_{-\infty}^{x} f(u)\,du$$

همچنین از خاصیت فوق وقتی $\Delta x \to 0$ داریم:

$$P\{x \le \mathbf{x} \le x + \Delta x\} = f(x)\Delta x$$ (مفهوم چگالی)

یا:

$$f(x) = \lim_{\Delta x \to 0} \frac{P\{x \le \mathbf{x} \le x + \Delta x\}}{\Delta x}$$



(این در واقع $f(x^+)$ است، چون CDF یک متغیر تصادفی پیوسته اگر چه پیوسته است، ولی لزوماً مشتق‌پذیر نیست. ممکن است مشتق راست و چپ برابر نباشند، در این صورت ناپیوستگی در $f(x)$ خواهیم داشت، مانند مثال بالا. در چنین نقاطی که مشتق وجود ندارد هر مقداری را می‌توانید به $f(x)$ نسبت دهید، مثلاً مشتق راست را.)

2) $f(x) \ge 0$

چون $F(x)$ غیرنزولی است.

(f چگالی احتمال مانند P تابع احتمال نامنفی است، ولی f چگالی احتمال می‌تواند بزرگتر از یک هم بشود و حتی می‌تواند به سمت بی‌نهایت برود، مثلاً $f(x) = \frac{2}{3}x^{-\frac{1}{3}}u(x)$)

$$3)\ \int_{-\infty}^{+\infty} f(x)\,dx = 1$$

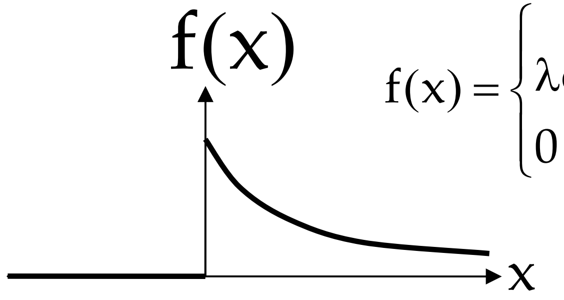
چون $F(+\infty) = 1$ است (مشابه خاصیت $\sum_{i=1}^{\infty} P(x_i) = 1$ برای تابع احتمال).

مثال: مدت زمان کارکرد یک کامپیوتر (بر حسب روز) قبل از خرابی یک متغیر تصادفی پیوسته با تابع چگالی زیر است:

الف) احتمال اینکه کامپیوتر بین ۵۰ تا ۱۵۰ روز کار کند چقدر است؟

ب) احتمال اینکه کمتر از ۱۰۰ روز کار کند چقدر است؟

$$f(x)=\begin{cases}\lambda e^{-\frac{x}{100}} & x\ge 0\\ 0 & x<0\end{cases}$$



الف)

$$\int_{-\infty}^{+\infty} f(x)\,dx=1\Rightarrow \lambda\int_0^{+\infty} e^{-\frac{x}{100}}\,dx=1\Rightarrow -\lambda(100)e^{-\frac{x}{100}}\Big|_0^{+\infty}=100\lambda=1\Rightarrow \lambda=\frac{1}{100}$$

$$P\{50<\mathbf{x}<150\}=\int_{50}^{150}\frac{1}{100}e^{-\frac{x}{100}}\,dx=-e^{-\frac{x}{100}}\Big|_{50}^{150}=e^{-\frac{1}{2}}-e^{-\frac{3}{2}}=0.384$$

ب)

$$P\{\mathbf{x}<100\}=\int_0^{100}\frac{1}{100}e^{-\frac{x}{100}}\,dx=-e^{-\frac{x}{100}}\Big|_0^{100}=1-e^{-1}=0.632$$

یعنی در ۶۳٪ اوقات کامپیوتر قبل از اینکه ۱۰۰ روز کار کند از کار می‌افتد.

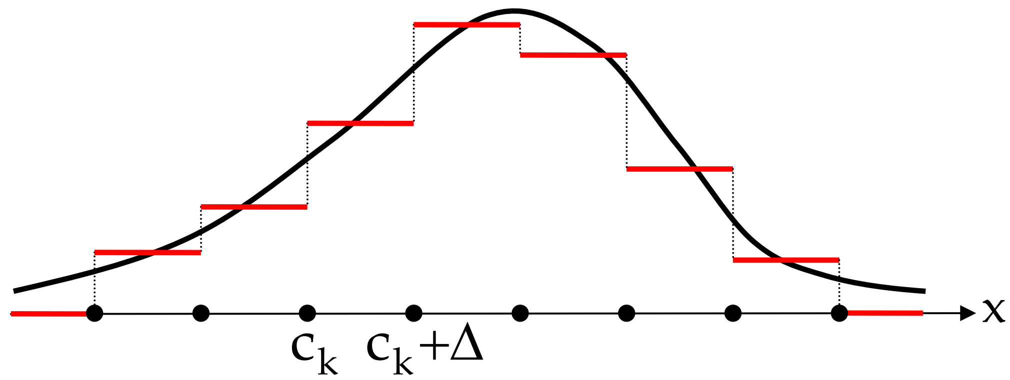
به دست آوردن تجربی $f(x)$:

آزمایش را nبار تکرار می‌کنیم و مقدار متغیر تصادفی $\mathbf{x}$ حاصله از هر آزمایش را ثبت می‌کنیم. محور xها را به فواصلی به طول Δ تقسیم می‌کنیم. اگر تعداد x_iهایی که در فاصلهٔ kام، یعنی $[c_k, c_k+\Delta)$ قرار دارند را با n_k نشان دهیم، داریم:

$$f(x) \simeq f_n(x) \text{ هیستوگرام } = \frac{n_k}{n\Delta} : c_k \le x < c_k + \Delta$$

به این ترتیب:

$$f(c_k)\Delta \simeq P\{c_k \le \mathbf{x} < c_k + \Delta\} \simeq \frac{n_k}{n} = f_n(x)\Delta$$



pdf برای متغیرهای تصادفی غیر پیوسته:

در مورد متغیر تصادفی غیرپیوسته یا مخلوط چون CDF دارای ناپیوستگی است، مشتق بی‌مفهوم خواهد بود (بی‌نهایت می شود). ولی با استفاده از تابع δ (تابع تعمیم یافتهٔ) می‌توانیم آن را بیان کنیم.

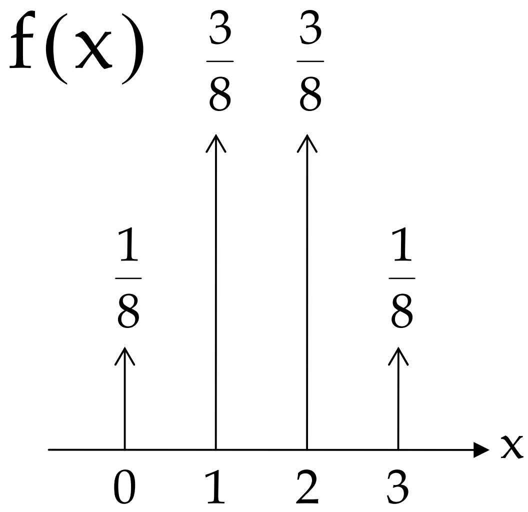
مثلاً برای متغیر تصادفی گسسته در مثال پرتاب سه سکه داریم:

$$f(x)=\sum_i P_i\,\delta(x-x_i)$$

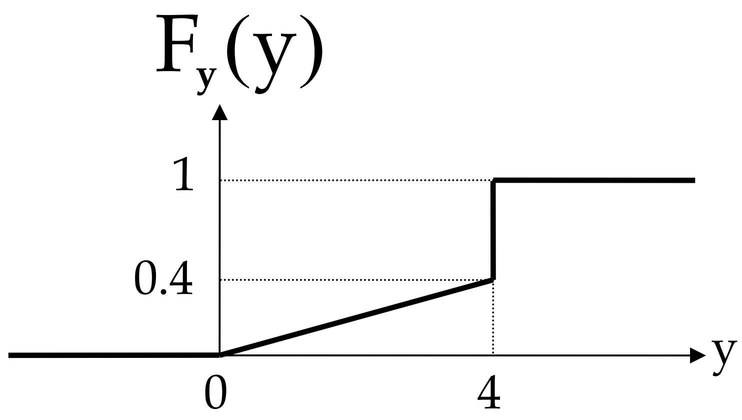
همان طور که داشتیم:

$$F(x)=\sum_i P_i\,u(x-x_i)$$

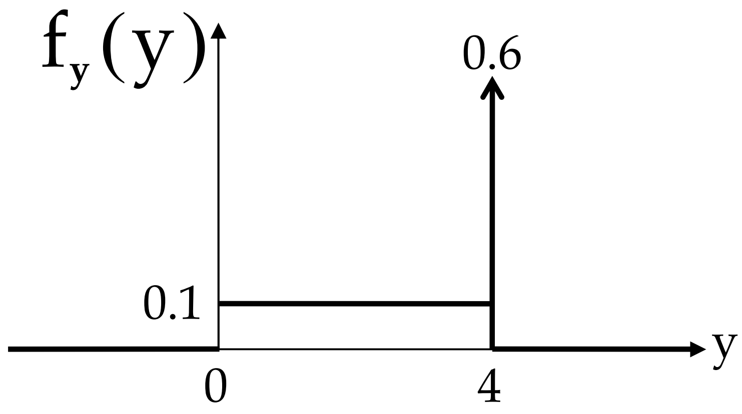
(یعنی نقاط تابع احتمال به دلتاهایی با آن وزنه‌ها تبدیل شد)



در مثال تقویت‌کننده‌ای که اشباع می شود نیز دیدیم:



لذا:



وقتی اجازه دهیم که $f(x)$ حاوی دلتا باشد، خواهیم داشت:

$$P\{x_1 \le \mathbf{x} \le x_2\} = \int_{x_1^-}^{x_2^+} f(x)\,dx$$

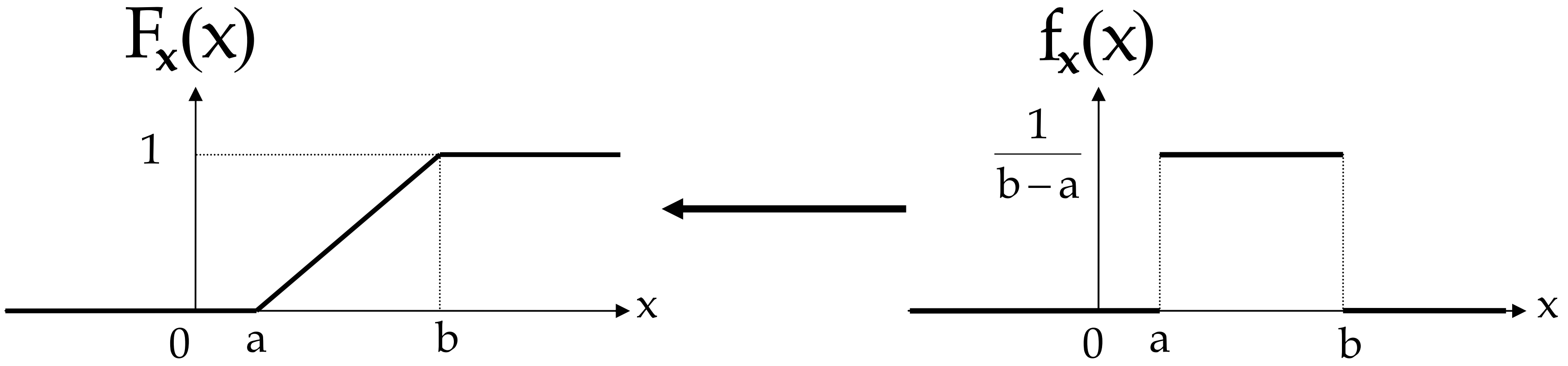
$$P\{x_1 < \mathbf{x} \le x_2\} = \int_{x_1^+}^{x_2^+} f(x)\,dx$$

برخی توزیع های خاص:

به سادگی می‌توان نشان داد که برای هر تابع غیرنزولی (و از راست پیوسته) که در بی‌نهایت، یک و در منهای بی‌نهایت صفر باشد، متغیر تصادفی‌ای وجود دارد که این تابع، تابع توزیع انباشتهٔ آن باشد (قضیهٔ وجود). ولی برخی توزیعها کاربرد بیشتری دارند.

توزیع یکنواخت (Uniform):

گوییم x دارای توزیع یکنواخت بین a و b است؛ $\mathbf{x} \sim u(a,b)$، اگر:



مثلاً اغلب فاز سیگنال دریافتی، یکنواخت بین 0 و 2π فرض می‌شود. مواقعی که هیچ اطلاع پیشینی در مورد نحوهٔ توزیع نداریم نیز با توجه به اصل ناکافی بودن دلیل از توزیع یکنواخت استفاده می‌شود.

توزیع نرمال (Normal) یا گوسی (Gaussian):

کاربرد بسیار زیادی دارد و در بسیاری موارد عملی تقریب خوبی برای pdf متغیر تصادفی مورد نظر عمل می‌کند. مثلاً در یک جامعهٔ همگن، توزیع قد و وزن افراد نرمال است. توزیع مشخصات قطعات تولیدی یک کارخانه معمولاً نرمال است. طبق **قضیهٔ حد مرکزی** مجموع متغیرهای تصادفی، ولو نرمال نباشند، به نرمال میل می‌کند. نشان داده شده که نویز حرارتی، ؟؟؟ نویز و ... گوسی هستند. از نظر ریاضی نیز کار با آن ساده است.

نرمال استاندارد؛ $\mathbf{x} \sim \mathrm{N}(0,1)$، یعنی همان $g(x)$ای که معرفی کردیم:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = g(x)$$

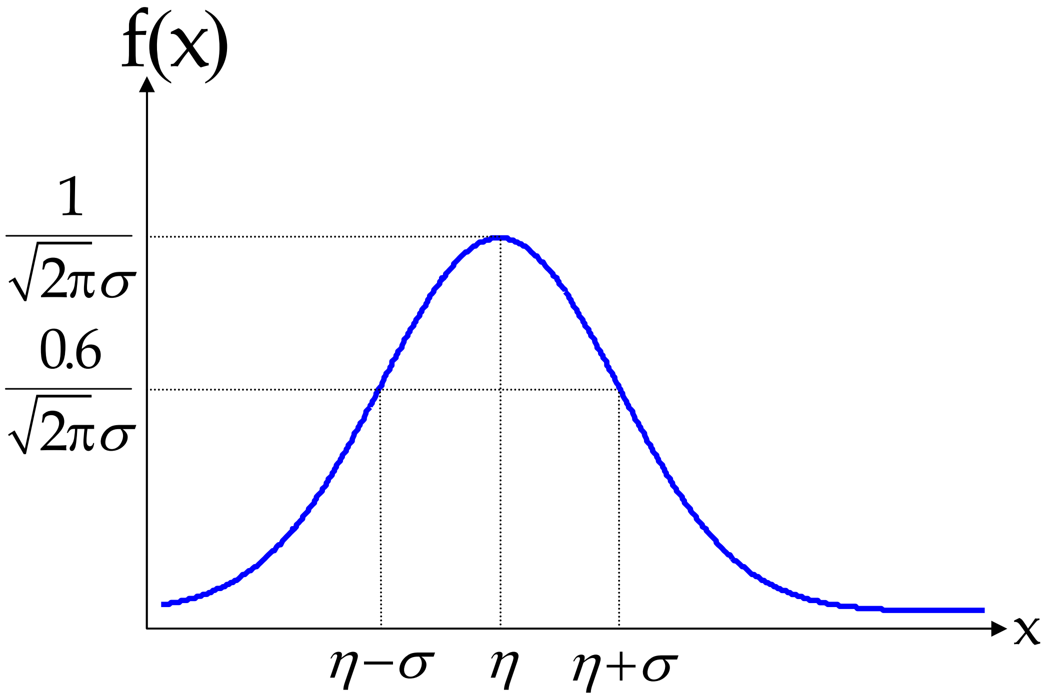
$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du = G(x)$$

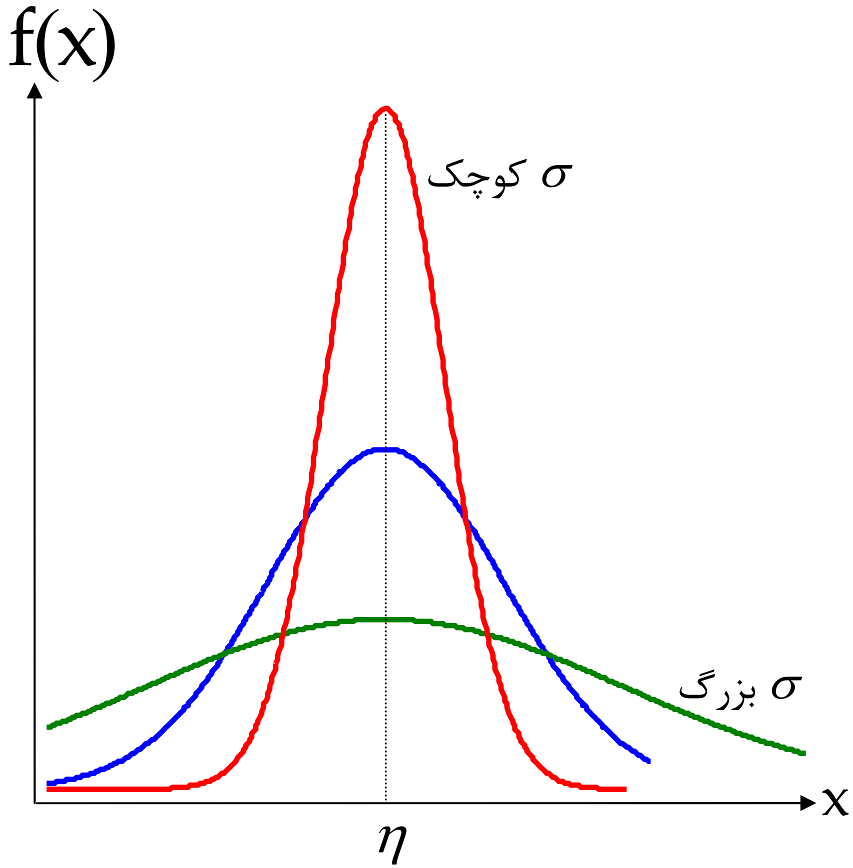
متغیر تصادفی نرمال استاندارد خیلی اوقات با $\mathbf{z}$ نشان داده می‌شود.

در حالت کلی (برای $\sigma > 0$)، اگر $\mathbf{x} \sim N(\eta,\sigma)$، داریم:

$$f_{\mathbf{x}}(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\eta)^2}{2\sigma^2}} = \frac{1}{\sigma} g\left(\frac{x-\eta}{\sigma}\right)$$

$$F_{\mathbf{x}}(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{x} e^{-\frac{(u-\eta)^2}{2\sigma^2}}\, du = G\left(\frac{x-\eta}{\sigma}\right)$$





اصولاً اگر $\mathbf{x} \sim N(\eta,\sigma)$ باشد، $a\mathbf{x}+b \sim N(a\eta+b,a\sigma)$ خواهد بود و لذا $\frac{\mathbf{x}-\eta}{\sigma} \sim N(0,1)$ می‌شود.

$$\mathbf{y} = a\mathbf{x} + b$$

$$F_{\mathbf{y}}(y) = P\{\mathbf{y} \le y\} = P\{a\mathbf{x}+b \le y\} = P\{\mathbf{x} \le \frac{y-b}{a}\} = \int_{-\infty}^{\frac{y-b}{a}} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\eta)^2}{2\sigma^2}}\, du$$

$$= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}\,a\sigma} e^{-\frac{[t-(a\eta+b)]^2}{2a^2\sigma^2}}\, dt = \int_{-\infty}^{y} f_{\mathbf{y}}(t)\,dt$$

$$P\{x_1 \le \mathbf{x} \le x_2\} = P\{\frac{x_1-\eta}{\sigma} \le \underbrace{\frac{\mathbf{x}-\eta}{\sigma}}_{\mathbf{z}} \le \frac{x_2-\eta}{\sigma}\} = F_{\mathbf{z}}(\frac{x_2-\eta}{\sigma}) - F_{\mathbf{z}}(\frac{x_1-\eta}{\sigma}) = G(\frac{x_2-\eta}{\sigma}) - G(\frac{x_1-\eta}{\sigma})$$

$$P\{\eta-\sigma \le \mathbf{x} \le \eta+\sigma\} \simeq 0.683$$
$$P\{\eta-2\sigma \le \mathbf{x} \le \eta+2\sigma\} \simeq 0.954$$
$$P\{\eta-3\sigma \le \mathbf{x} \le \eta+3\sigma\} \simeq 0.997$$

خوب است به خاطر داشته باشید که در توزیع نرمال، در خیلی از کتابها $G(x)$ با $\Phi(x)$ نشان داده می‌شود:

Kreyszig, Tables A8 and A9.

در کاربردهای مهندسی برق خیلی اوقات $F^c(x) = P\{\mathbf{x} > x\}$ نرمال را لازم داریم:

$$Q(x) = 1 - G(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

بعضاً نیز بر حسب تابع معروف Error Function (تابع خطا) بیان می‌کنند:

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_{0}^{x} e^{-u^2} du$$

$$G(x) = \frac{1}{2} + \frac{1}{2} erf\left(\frac{x}{\sqrt{2}}\right)$$

مثال: ولتاژ نویزی دارای توزیع نرمال با $\eta = 0$ و $\sigma = 0.75$ است. اگر **x** نشان دهندهٔ ولتاژ نویز باشد؛

الف) احتمال اینکه $|\mathbf{x}| \le 1.5\,\mathrm{Volts}$ باشد چقدر است؟

ب) ولتاژ نویز **x** از چه سطح ولتاژی در ۹۹٪ موارد تجاوز نمی‌کند؟

ج) $|\mathbf{x}|$ از چه سطح ولتاژی در ۹۹٪ موارد تجاوز نمی‌کند؟

الف)

$$\mathrm{P}\{|\mathbf{x}| \le 1.5\} = \mathrm{P}\{-1.5 \le \mathbf{x} \le 1.5\} = G(\frac{1.5-0}{0.75}) - G(\frac{-1.5-0}{0.75}) = 2G(\frac{1.5}{0.75}) - 1 = 0.954$$

چون این همان محدودهٔ $\eta \pm 2\sigma$ است، از قبل هم می‌دانستیم ۹۵٪ می شود.

ب)

$$\mathrm{P}\{\mathbf{x} \le \mathrm{x}\} = 0.99 \rightarrow$$ یعنی دنبال صدک ۹۹ام هستیم

$$G(\frac{\mathrm{x}-0}{0.75}) = 0.99$$

از جدول A9 می‌توانیم G^{-1} را حساب کنیم:

$$\frac{\mathrm{x}-0}{0.75} = G^{-1}(0.99) = 2.326 \Rightarrow \mathrm{x} = 1.745$$: صدک نود و نهم

ج)

$$P\{|\mathbf{x}| \le x\} = 0.99$$

$$2G\left(\frac{x-0}{0.75}\right) - 1 = 0.99 \Rightarrow \frac{x}{0.75} = G^{-1}(0.995) \Rightarrow x = 1.932$$

یا از جدول A9 مستقیماً D^{-1} را داریم:

$$D(z) = G(z) - G(-z)$$

$$\frac{x-0}{0.75} = 2.576 \Rightarrow x = 1.932$$

توزیع نمایی (Exponential):

$$f_x(x) = \lambda e^{-\lambda x} : x \geq 0$$

یا می‌نویسیم:

$$f_x(x) = \lambda e^{-\lambda x} u(x) \Rightarrow F_x(x) = (1 - e^{-\lambda x}) u(x)$$

این توزیع کاربرد بسیار زیادی دارد. مثلاً فاصلهٔ بین دو نقطهٔ تصادفی دارای توزیع نمایی است.

مکالمات تلفنی:

تعداد این نقاط در یک بازهٔ زمانی مشخص دارای توزیع پواسن است، اما فاصلهٔ بین آنها توزیع نمایی دارد.

t می‌تواند مقادیر بین صفر تا بی‌نهایت را اختیار کند.

$$f_{\mathbf{t}}(\tau) d\tau = P\{\tau < \mathbf{t} \leq \tau + d\tau\}$$

t وقتی بین τ و $\tau+d\tau$ خواهد بود که تا قبل از τ مکالمه‌ای اتفاق نیفتاده باشد و بین τ و $\tau+d\tau$ یک مکالمه اتفاق افتد.

$$P\{\tau \text{ در } k\text{نقطه}\} = e^{-\lambda \tau} \frac{(\lambda \tau)^k}{k!}$$

$$f_t(\tau)d\tau = P\{\tau \text{ عدم وقوع در فاصلهٔ }\}P\{d\tau \text{ یک بار وقوع در }\} = e^{-\lambda\tau}\frac{(\lambda\tau)^0}{0!}e^{-\lambda d\tau}\frac{(\lambda d\tau)^1}{1!}$$

$$f_t(\tau)d\tau = \lambda e^{-\lambda(\tau+d\tau)}d\tau$$

$$d\tau \to 0 \Rightarrow f_t(\tau) = \lambda e^{-\lambda\tau} : \tau \geq 0$$

یا به عبارت بهتر:

$$F_t(\tau) = P\{t \leq \tau\} = P\{\tau \text{ لااقل یک نقطه در فاصلهٔ }\} = P\{k \geq 1\} = 1 - P\{k = 0\} = 1 - e^{-\lambda\tau}\frac{(\lambda\tau)^0}{0!} = 1 - e^{-\lambda\tau} : \tau \geq 0$$

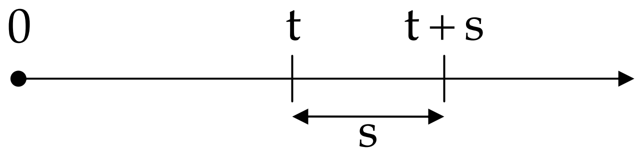
فاصلهٔ بین دو زمین‌لرزهٔ متوالی و مستقل، فاصلهٔ بین دو بار خراب شدن یک دستگاه و ... (با فرض استقلال از هم و متساوی الاحتمال بودن در تمام زمانها) توزیع نمایی دارند.

لذا این توزیع برای تصادف اتومبیل و طوفانهای خورشیدی قابل استفاده نیست.

دیده شد که طول مکالمهٔ تلفنی هم توزیع نمایی دارد.

از خواص جالب توزیع نمایی این است که بدون حافظه است، یعنی:

$$P\{\mathbf{x} > t+s \mid \mathbf{x} > t\} = P\{\mathbf{x} > s\}$$



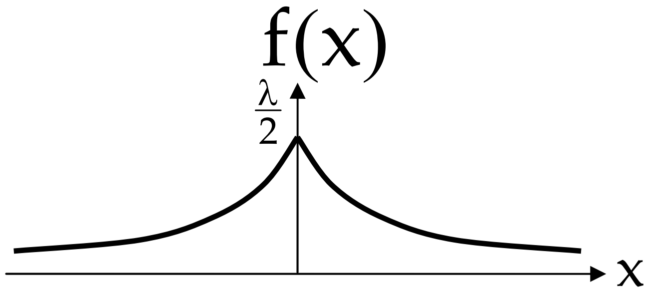
زیرا:

$$P\{\mathbf{x} > t+s \mid \mathbf{x} > t\} = \frac{1-F_{\mathbf{x}}(t+s)}{1-F_{\mathbf{x}}(t)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s} = 1-F_{\mathbf{x}}(s) = P\{\mathbf{x} > s\}$$

می توان نشان داد که تنها تابع توزیعی که دارای این خاصیت می‌باشد تابع نمایی است.

توزیع لاپلاس (Laplace):

$$f(x) = \frac{\lambda}{2} e^{-\lambda|x|}$$



این توزیع در مورد نویزهای اسپایکی مدل خوبی است.

توزیع گاما (Gamma):

$$f(x) = A\,x^{r-1}\,e^{-\lambda x}\,u(x) : r > 0, \lambda > 0$$

از این توزیع نیز درکاربردهای مختلفی استفاده می‌شود.

ضریب ثابت A را باید طوری به دست آوریم که:

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_{0}^{+\infty} A x^{r-1} e^{-\lambda x}\,dx = 1$$

$$y = \lambda x \Rightarrow \frac{A}{\lambda^{r}} \int_{0}^{+\infty} y^{r-1} e^{-y}\,dy = 1$$

از طرفی می‌دانیم که **تابع گاما** برابر است با:

$$\Gamma(r) = \int_{0}^{+\infty} y^{r-1} e^{-y}\,dy$$

در واقع $\Gamma(\bullet)$، فاکتوریل تعمیم یافته است. چون $\Gamma(r) = (r-1)\Gamma(r-1)$، برای r صحیح و مثبت داریم: $\Gamma(r) = (r-1)!$.

$$\frac{A}{\lambda^r}\Gamma(r)=1 \Rightarrow A=\frac{\lambda^r}{\Gamma(r)}$$

$$f_{\mathbf{x}}(x)=\frac{\lambda(\lambda x)^{r-1}e^{-\lambda x}}{\Gamma(r)}u(x)$$

برای نشان دادن این توزیع می‌نویسیم $\mathbf{x} \sim \text{Gamma}(r,\lambda)$ است.

شکل صفحهٔ ۸۵ کتاب Helstrom برای $\lambda=1$ و rهای مختلف.

حالت خاص: اگر $r=1$ باشد، داریم:

توزیع نمایی: $f(x)=\lambda e^{-\lambda x}u(x)$

حالت خاص: اگر $\lambda=\frac{1}{2}$ و $r=\frac{n}{2}$ که n عددی صحیح است، داریم:

$$f(x)=\frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}x^{(\frac{n}{2}-1)}e^{-\frac{x}{2}}u(x)$$

حالت خاص: اگر $r = n$ که n عددی صحیح است، داریم:

$$f(x) = \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x} u(x)$$

این توزیع مدت زمان لازم برای وقوع n واقعهٔ کاملاً تصادفی (حاصل جمع n نمایی مستقل از هم) است که **توزیع ارلانگ (Erlang)** نام دارد و در تئوری صف و ترافیک (مکالمات تلفنی)، تشعشع رادیواکتیو و ... استفاده می‌شود.

توزیع χ^2 (Chi-Square) با n درجه آزادی (n degrees of freedom):

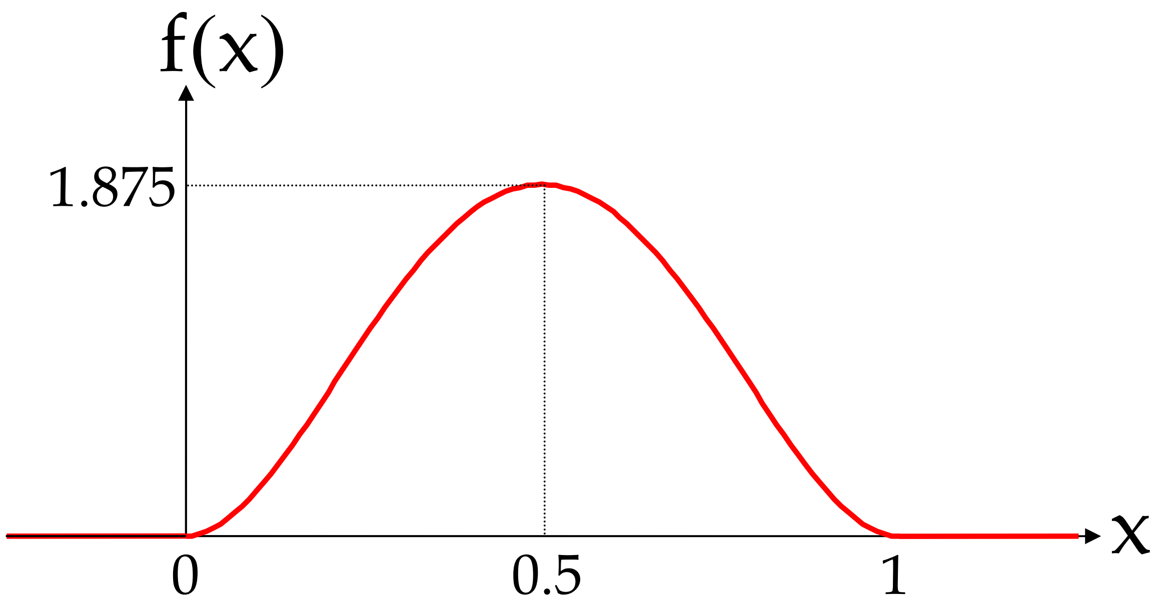
اگر $\mathbf{y} = \mathbf{x}_1^2 + \mathbf{x}_2^2 + \cdots + \mathbf{x}_n^2$ که $\mathbf{x}_i$ها $N(0,1)$ بوده و مستقل باشند، $\mathbf{y}$ دارای توزیع $\chi^2(n)$ خواهد بود. این توزیع در مهندسی برق و آمار استفادهٔ زیادی دارد.

توزیع بتا (Beta):

$$f(x)=\begin{cases}\dfrac{1}{\beta(a,b)}x^{a-1}(1-x)^{b-1} & 0\le x\le 1\\ 0 & \text{otherwise}\end{cases}$$

اگر $a=b=1$ باشد، همان توزیع یکنواخت روی $[0,1]$ خواهد بود.

برای $a=b=3$، نمودار توزیع به این صورت خواهد بود:



یادآوری: تابع بتا برابر است با:

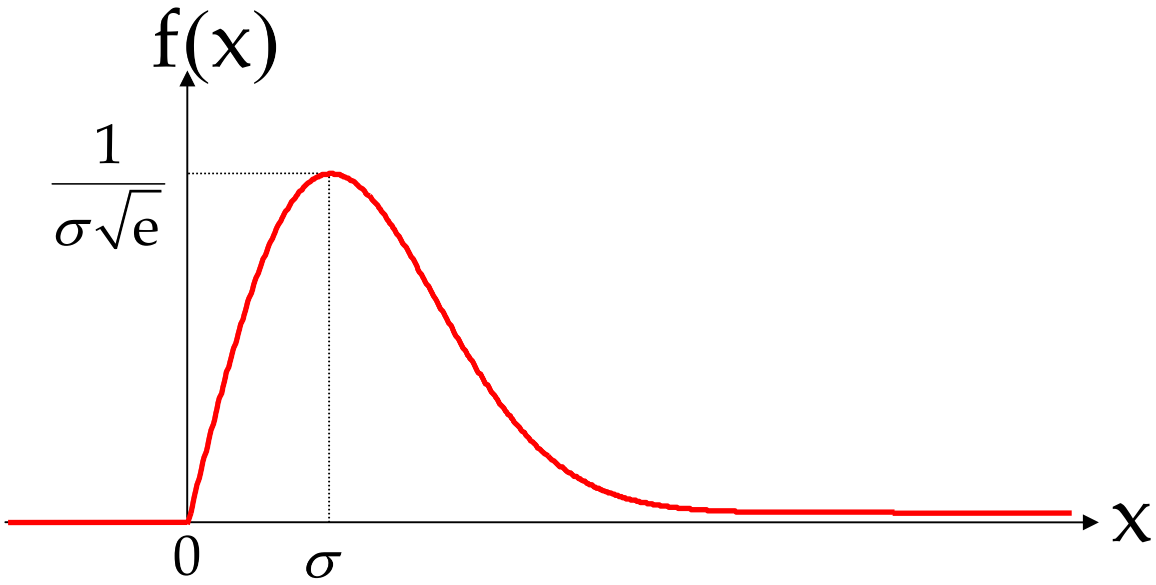
$$\beta(a,b)=\int_0^1 x^{a-1}(1-x)^{b-1}dx=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

پس:

$$f(x)=\begin{cases}\dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}x^{a-1}(1-x)^{b-1} & 0\le x\le 1\\ 0 & \text{otherwise}\end{cases}$$

توزیع رایلی (Rayleigh):

$$f(x)=\frac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}u(x)$$

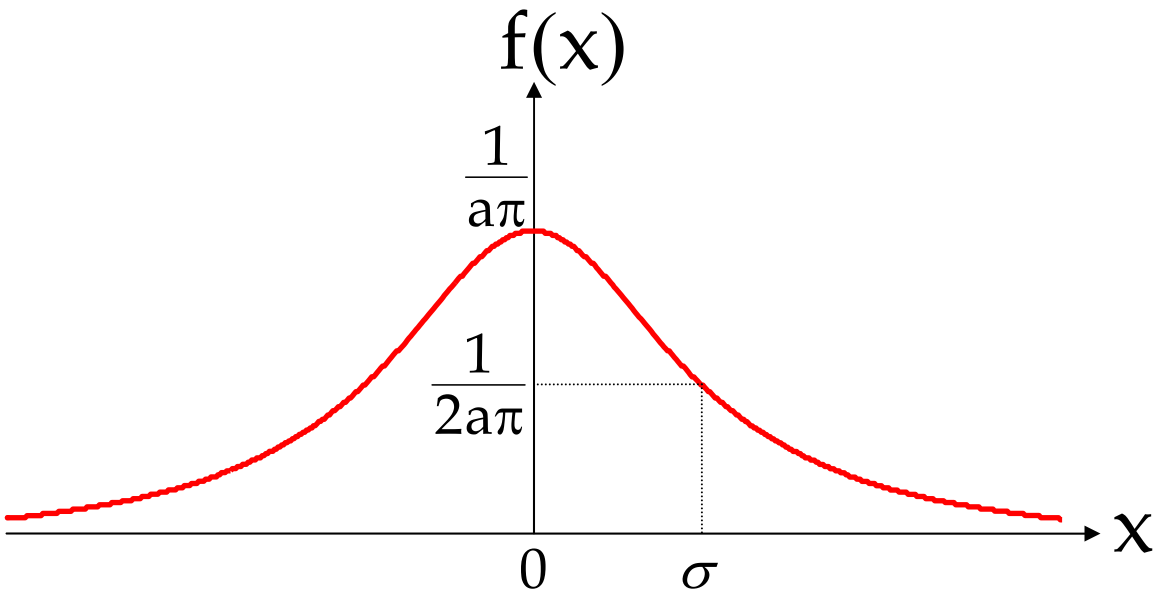


اگر فرض کنیم $y=\sqrt{x_1^2+x_2^2}$ باشد و x_1 و x_2 مستقل و دارای توزیع نرمال $N(0,\sigma)$ باشند، y دارای توزیع رایلی خواهد بود.
مثلاً اگر $z=x+jy$ باشد، $|z|$ دارای توزیع رایلی خواهد بود.

این توزیع نیز کاربرد بسیار زیادی در مهندسی برق دارد، مانند پوش نویز گوسی.

توزیع کوشی (Cauchy):

$$f(x)=\frac{\frac{a}{\pi}}{x^2+a^2}$$



اگر θ دارای توزیع یکنواخت $[-\frac{\pi}{2},\frac{\pi}{2}]$ باشد، $a\,tg\theta$ دارای این توزیع خواهد بود.

توزیع دوجمله‌ای:

این یک توزیع گسسته است. اگر در یک آزمایش تصادفی احتمال وقوع واقعهٔ A برابر $P(A)=p$ باشد و برای آزمایش حاصل از nبار تکرار این آزمایش، متغیر تصادفی **x** را به این صورت تعریف کنیم که: **x** = تعداد وقوع واقعهٔ A در n آزمایش، **x** دارای توزیع دوجمله‌ای خواهد بود:

$$\mathbf{x} \sim \text{Binomial}(n,p)$$

$$P(k)=P\{\mathbf{x}=k\}=\binom{n}{k}p^k q^{n-k} : k=0,1,2,\ldots,n$$

و برای سایر مقادیر x داریم: $P(x)=0$.

مثال: در جعبه‌ای N دیود قرار دارند که $K\le N$تا از آنها خرابند. N نمونه با جایگزینی برداشته می‌شود. اگر متغیر تصادفی **x** به صورت: **x** = تعداد خرابها، تعریف شود، $P(x)$ را به دست آورید.

$$P(x)=\begin{cases}\binom{n}{x}\left(\frac{K}{N}\right)^x\left(1-\frac{K}{N}\right)^{n-x} & x=0,1,2,\ldots,n \\ 0 & \text{otherwise}\end{cases}$$

حالت خاص: اگر $n=1$ باشد، یعنی آزمایش تصادفی را فقط یک بار انجام می‌دهیم. پس **x** فقط می‌تواند صفر یا یک باشد (صفر با احتمال q و یک با احتمال p)، پس داریم:

$$P(k)=p^k q^{1-k} : k=0,1$$

یعنی $P(0)=q$ و $P(1)=p$ می‌باشد.

این حالت خاص را **توزیع برنولی** گویند.

با توجه به قضیهٔ لاپلاس، برای n بزرگ توزیع (گسستهٔ) دوجمله‌ای را می‌توان با توزیع (پیوستهٔ) نرمال تقریب زد:

$$P(x)=\binom{n}{k}p^k q^{n-k}\simeq\frac{1}{\sqrt{2\pi npq}}e^{-\frac{(k-np)^2}{2npq}}=N(np,\sqrt{npq})$$

$$F(x)\simeq G(\frac{x-np}{\sqrt{npq}})$$

و با توجه به قضیهٔ پواسون، برای n بزرگ و p کوچک می‌توان توسط توزیع پواسون تقریب زد:

$$\binom{n}{k}p^k q^{n-k}\simeq e^{-np}\frac{(np)^k}{k!}=\text{Poisson}(np)$$

توزیع فوق‌هندسی (Hypergeometric):

در مثال قبل اگر انتخاب بدون جایگزینی باشد، داریم:

$$P(x)=\frac{\text{تعداد انتخابهای مورد نظر}}{\text{تعداد کل انتخابها}}=\begin{cases}\dfrac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}} & \max(0,n-N+K)\le k\le \min(n,K)\\ 0 & \text{otherwise}\end{cases}$$

$\max(0,n-N+K)$ بدین خاطر است که تعداد سالم‌های انتخاب‌شده n-k بوده و نمی‌تواند از تعداد کل سالم‌ها که N-K است، بیشتر شود، پس: $n-k\le N-K$؛

$\min(n,K)$ نیز بدین خاطر است که تعداد خرابهای انتخاب‌شده نمی‌تواند از تعداد کل خرابها بیشتر شود.

یا اگر بگیریم: $p=\frac{K}{N}$، خواهیم داشت:

$$P(k)=\frac{\binom{Np}{k}\binom{N-Np}{n-k}}{\binom{N}{n}}$$

توزیع دوجمله‌ای منفی (پاسکال):

فرض کنید با مسئلهٔ زیر مواجهیم. سکه‌ای را آنقدر پرتاب می‌کنیم که rبار شیر بیاید. اگر تعداد کل پرتابها را $\mathbf{y}$ بنامیم، احتمال $\mathbf{y}=n$ را به دست آورید $(n=r,r+1,\ldots)$.

در حالت کلی اگر یک آزمایش برنولی را آنقدر تکرار کنیم تا واقعهٔ A که $P(A)=p$، rبار اتفاق افتد و متغیر تصادفی $\mathbf{y}$ به صورت:
$\mathbf{y}$ = تعداد کل آزمایشهای انجام شده تا rبار وقوع A تعریف شود، داریم:

احتمال rامین موفقیت که حتماً باید در nامین آزمایش باشد

$$P\{\mathbf{y}=n\}=p\underbrace{\binom{n-1}{k-1}p^{r-1}q^{(n-1)-(r-1)}}_{\text{احتمال } r-1 \text{ موفقیت در میان بقیهٔ آزمایشها (} n-1 \text{ آزمایش)}}$$

یعنی:

$$P\{\mathbf{y}=n\}=\binom{n-1}{r-1}p^{r}q^{n-r}\;:n=r,r+1,r+2,\ldots$$

در بعضی کتابها این را توزیع دوجمله‌ای منفی می‌گویند.

اما اگر فرض کنیم: $\mathbf{y}=\mathbf{x}+r$، که $\mathbf{x}$ تعداد شکستها و $\mathbf{y}$ تعداد کل آزمایشها است، واقعهٔ $\mathbf{x}=k$ با واقعهٔ $\mathbf{y}=k+r$ معادل بوده و احتمال یکسانی دارد، یعنی:

$$P\{\mathbf{x}=k\}=\binom{n-1}{r-1}p^r q^{n-r} \overset{n=k+r}{=} \binom{r+k-1}{r-1}p^r q^k=\binom{r+k-1}{k}p^r q^k : k=0,1,2,\dots$$

در تمرین‌ها داشتیم که:

$$\binom{r+k-1}{k}=(-1)^k\binom{-r}{k}$$

لذا:

$$P\{\mathbf{x}=k\}=P_{\mathbf{x}}(k)=\binom{-r}{k}p^r(-q)^k : k=0,1,2,\dots$$

این را **توزیع دوجمله‌ای منفی** گویند.

توزیع هندسی (Geometric):

حالت خاص توزیع دوجمله‌ای منفی است که در آن $r=1$ باشد، یعنی:

$$P\{\mathbf{x}=k\}=P_{\mathbf{x}}(k)=pq^{k} : k=0,1,2,\ldots$$

در اینجا، $\mathbf{x}$، تعداد شکستها در تکرار آزمایش تا حصول اولین موفقیت است.

این توزیع در واقع یک دنبالهٔ هندسی است و داریم:

$$\sum_{k=0}^{+\infty} pq^{k}=\frac{p}{1-q}=1$$

در بعضی کتابها حالت $r=1$ از آنچه ما با $P(\mathbf{y})$ نشان دادیم را توزیع هندسی گویند. برای $r=1$ داریم:

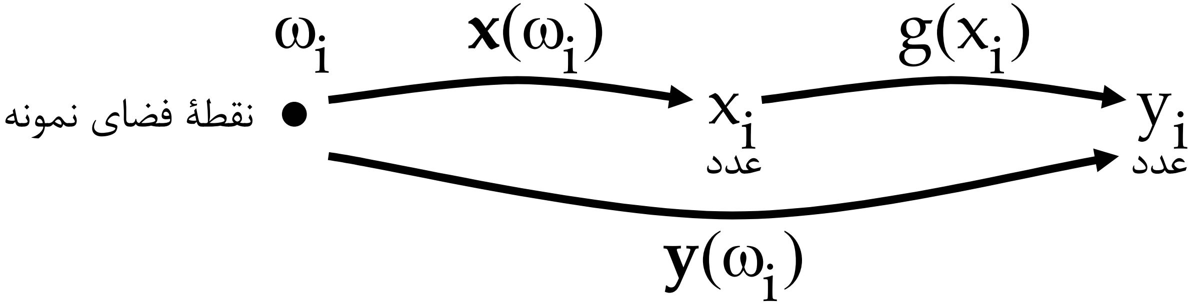
$$P\{\mathbf{y}=n\}=pq^{n-1} : n=1,2,3,\ldots$$

در اینجا، $\mathbf{y}$، تعداد کل آزمایشها تا حصول اولین موفقیت است.

تابع یک متغیر تصادفی:

در بسیاری موارد با فرض دانستن تابع توزیع یک متغیر تصادفی، با توجه به اینکه در سیستم مورد نظر عملیاتی روی آن انجام می‌شود، علاقه‌مندیم که توزیع تابعی از یک متغیر تصادفی را به دست آوریم. مانند آشکارساز دیودی. یا مثلاً مایلیم بدانیم که توزیع توان مصرفی مقاومتی که ولتاژ دو سر آن تصادفی است، چیست. یا اگر فاز سیگنال تصادفی باشد و $\mathbf{y} = A\cos(\omega t + \varphi)$، توزیع $\mathbf{y}$ چیست و امثالهم...

اگر $g(x)$ یک تابع حقیقی باشد، می‌توانیم به هر $x_i = \mathbf{x}(\omega_i)$، $y_i = g(x_i)$ را نسبت دهیم:



پس با ترکیب دو تابع مواجهیم که یک متغیر تصادفی جدید به ما می‌دهد:

$$\mathbf{y}(\omega) = g(\mathbf{x}(\omega))$$

متغیر تصادفی $\mathbf{y}$ را تابعی از متغیر تصادفی $\mathbf{x}$ گوییم.

حال ببینیم با دانستن توزیع $\mathbf{x}$ چگونه می‌توانیم توزیع $\mathbf{y}$ را به دست آوریم.

در حالت گسسته مسئله بسیار ساده است:

$$P_{\mathbf{y}}(y) = P\{\mathbf{y} = y\} = P\{g(\mathbf{x}) = y\} = \sum_{i:g(x_i)=y} P(x_i)$$

$$F_{\mathbf{y}}(y) = P\{\mathbf{y} \le y\} = P\{g(\mathbf{x}) \le y\} = \sum_{i:g(x_i)\le y} P(x_i)$$

مثلاً اگر متغیر تصادفی **x** مقادیر زیر را اختیار کند:

x_i	0	1	-1
P_i	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

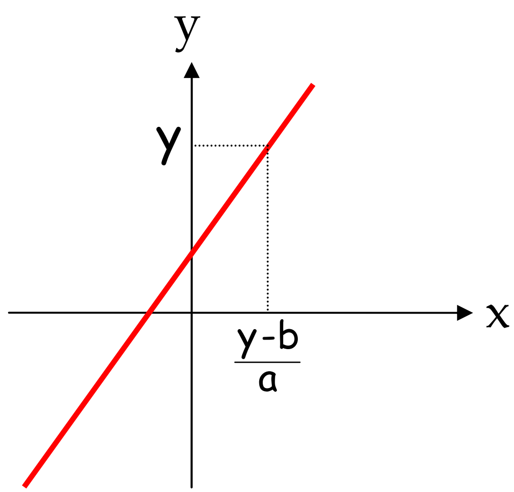
برای $\mathbf{y} = \mathbf{x}^2$ داریم:

$$P_{\mathbf{y}}(1) = P\{\mathbf{y} = 1\} = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

و به سادگی می‌توانیم بقیهٔ احتمالها را نیز حساب کنیم.

اما برای حالتهای غیرگسسته ابتدا به محاسبهٔ تابع CDF می‌پردازیم که ساده‌تر است.

مثال ۱: $g(x) = ax + b$:



$$\mathbf{y} = a\mathbf{x} + b$$

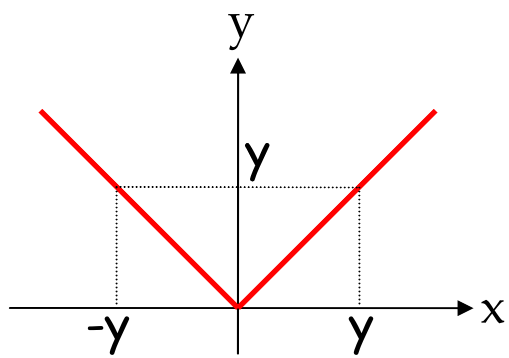
یعنی شیفت و scale

$$F_{\mathbf{y}}(y) = P\{\mathbf{y} \le y\} = P\{a\mathbf{x} + b \le y\} = \begin{cases} P\{\mathbf{x} \le \frac{y-b}{a}\} = \overbrace{F_{\mathbf{x}}(\frac{y-b}{a})} & a > 0 \\ P\{\mathbf{x} \ge \frac{y-b}{a}\} = 1 - F_{\mathbf{x}}(\frac{y-b}{a}) & a < 0 \end{cases}$$

از اینجا داریم:

$$f_{\mathbf{y}}(y) = \begin{cases} \frac{1}{a} f_{\mathbf{x}}(\frac{y-b}{a}) & a > 0 \\ -\frac{1}{a} f_{\mathbf{x}}(\frac{y-b}{a}) & a < 0 \end{cases} = \frac{1}{|a|} f_{\mathbf{x}}(\frac{y-b}{a})$$

مثال ۲ (یکسوساز تمام موج): $g(x) = |x|$:



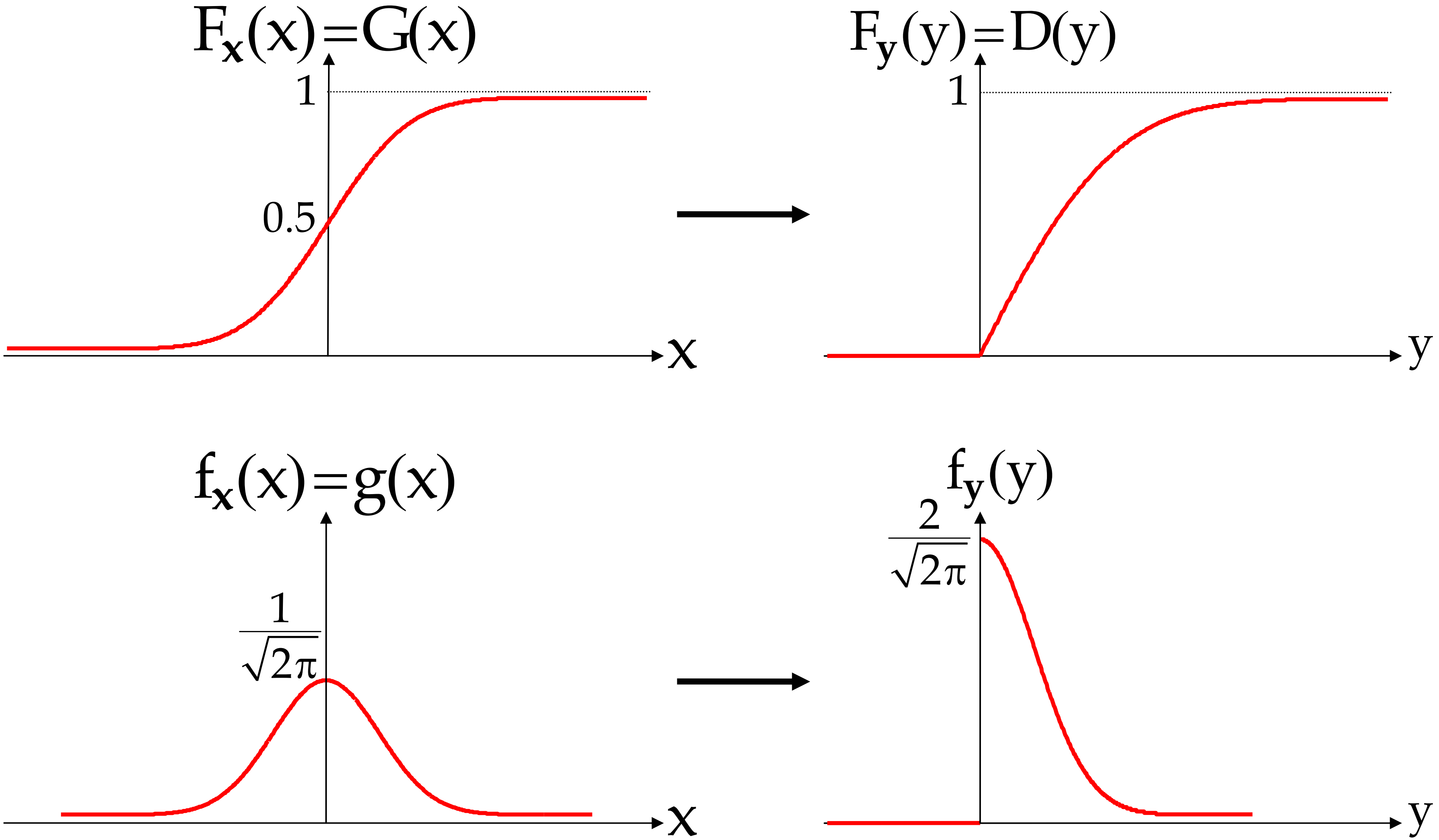
$$\mathbf{y} = |\mathbf{x}|$$

$$F_{\mathbf{y}}(y) = P\{\mathbf{y} \le y\} = P\{|\mathbf{x}| \le y\} = \begin{cases} 0 & y < 0 \\ P\{-y \le \mathbf{x} \le y\} = F_{\mathbf{x}}(y) - F_{\mathbf{x}}(-y) & y \ge 0 \end{cases}$$

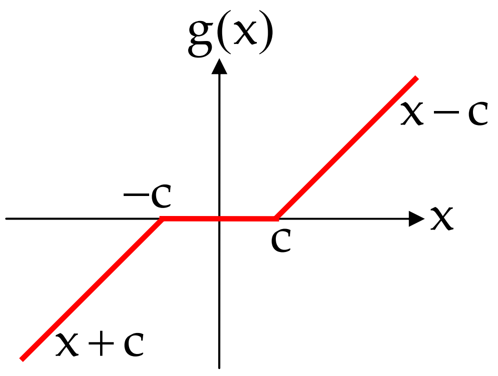
با مشتق‌گیری داریم:

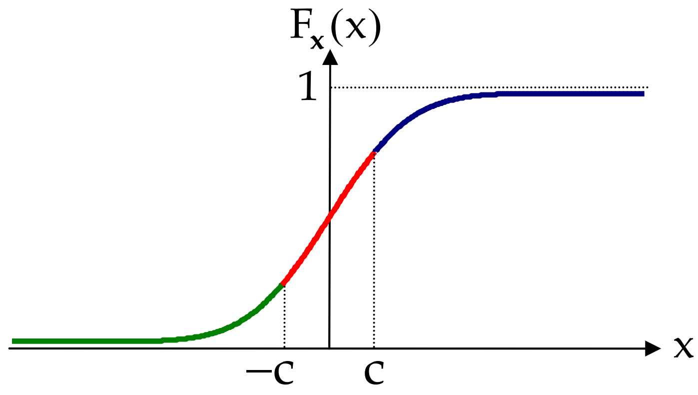
$$f_{\mathbf{y}}(y) = \begin{cases} f_{\mathbf{x}}(y) + f_{\mathbf{x}}(-y) & y \ge 0 \\ 0 & y < 0 \end{cases}$$

مثلاً اگر $\mathbf{x} \sim N(0,1)$ باشد، خواهیم داشت:



مثال ۳: $g(x)=\begin{cases} x+c & x<-c \\ 0 & -c\le x\le c \\ x-c & x>c \end{cases}$:

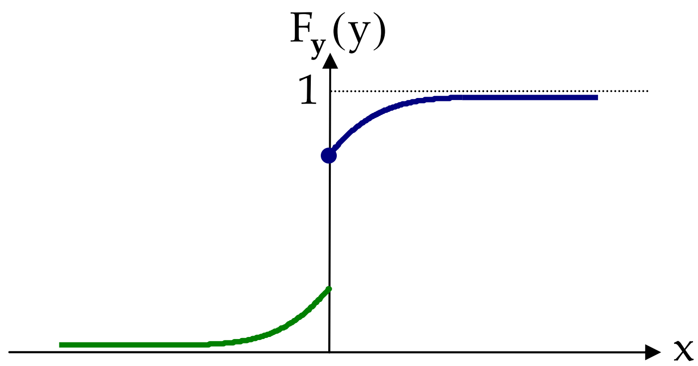




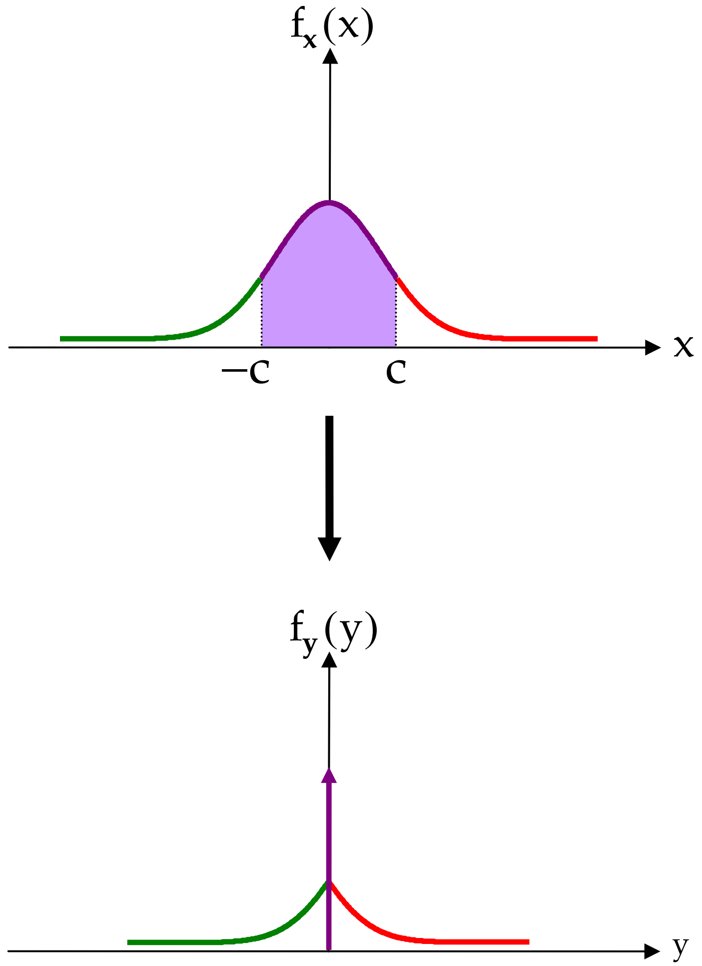
$$\mathbf{y}=g(\mathbf{x})$$

$$F_{\mathbf{y}}(y)=P\{\mathbf{y}\le y\}=\begin{cases} P\{\mathbf{x}+c\le y\} & y<0 \\ P\{\mathbf{x}-c\le y\} & y\ge 0 \end{cases}=\begin{cases} F_{\mathbf{x}}(y-c) & y<0 \\ F_{\mathbf{x}}(y+c) & y\ge 0 \end{cases}$$

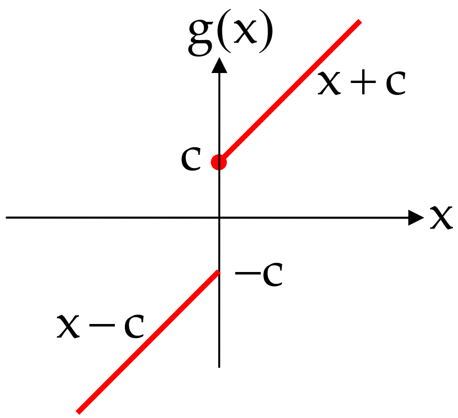
$$P\{\mathbf{y}=0\}=F_{\mathbf{x}}(c)-F_{\mathbf{x}}(-c)$$



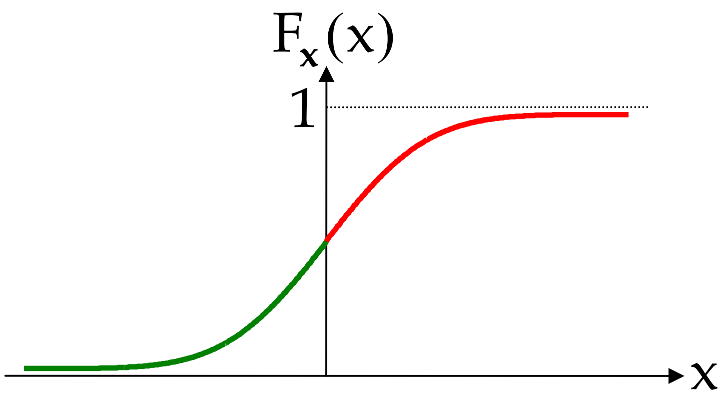
$$f_{\mathbf{y}}(y)=\begin{cases} f_{\mathbf{x}}(y-c) & y<0 \\ \delta(y)\int_{-c}^{+c} f_{\mathbf{x}}(u)du & y=0 \\ f_{\mathbf{x}}(y+c) & y>0 \end{cases}$$

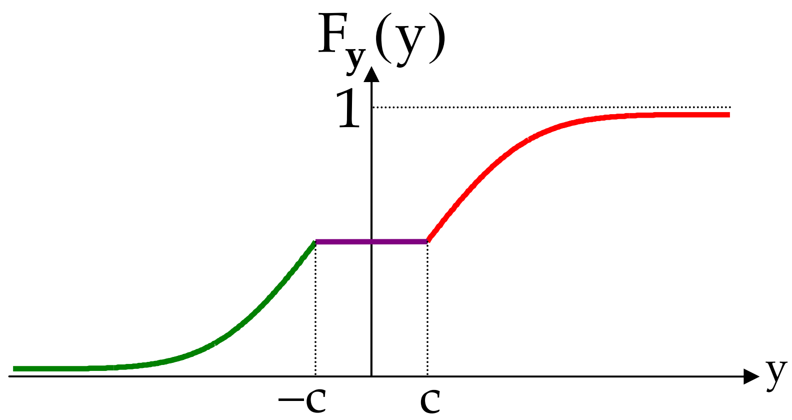


مثال ۴: $g(x)=\begin{cases} x+c & x\ge 0 \\ x-c & x<0 \end{cases}$:

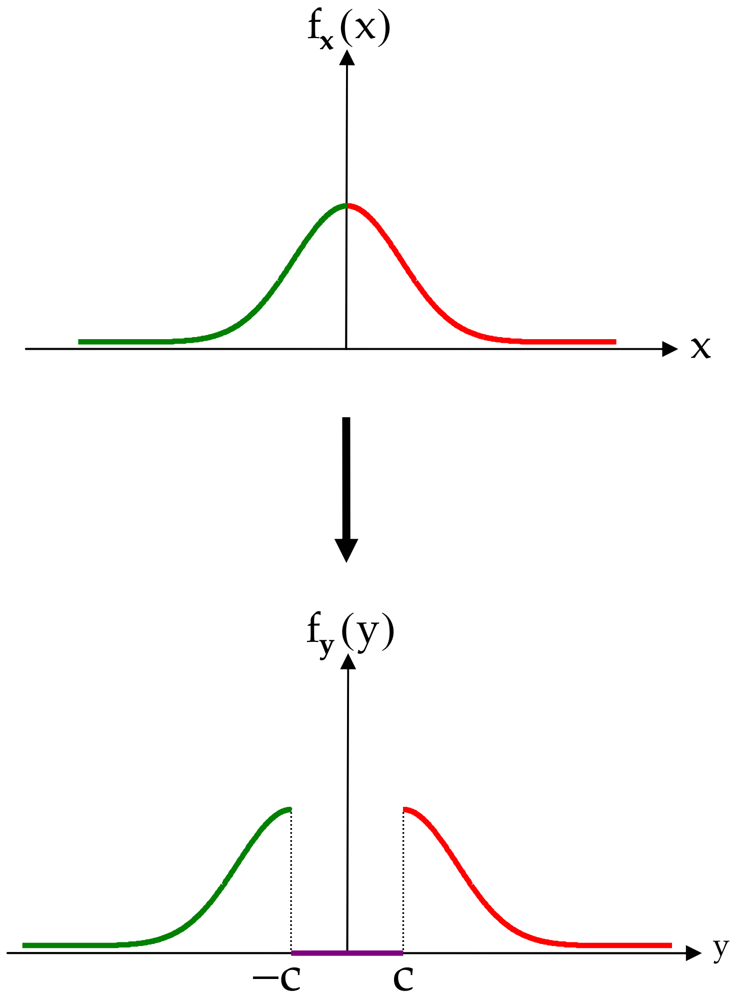


$$F_{\mathbf{y}}(y)=P\{\mathbf{y}\le y\}=\begin{cases} P\{\mathbf{x}+c\le y\} & y\ge c \\ P\{\mathbf{x}\le 0\} & -c<y<c \\ P\{\mathbf{x}-c\le y\} & y<-c \end{cases}=\begin{cases} F_{\mathbf{x}}(y-c) & y\ge c \\ F_{\mathbf{x}}(0) & -c<y<c \\ F_{\mathbf{x}}(y+c) & y<-c \end{cases}$$





$$f_y(y)=\begin{cases} f_x(y-c) & y \ge c \\ 0 & -c < y < c \\ f_x(y+c) & y < -c \end{cases}$$



به یاد داشته باشید که بی‌تغییر ماندن $g(x)$ در یک محدوده سبب پرش در F_y می‌شود و به عکس پرش در $g(x)$ موجب بی‌تغییر ماندن F_y در یک محدوده می‌شود.