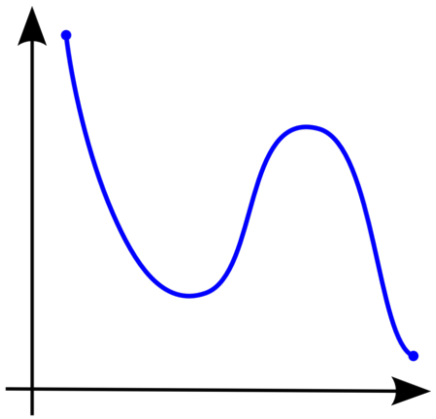
CS 215
Data Analysis and Interpretation

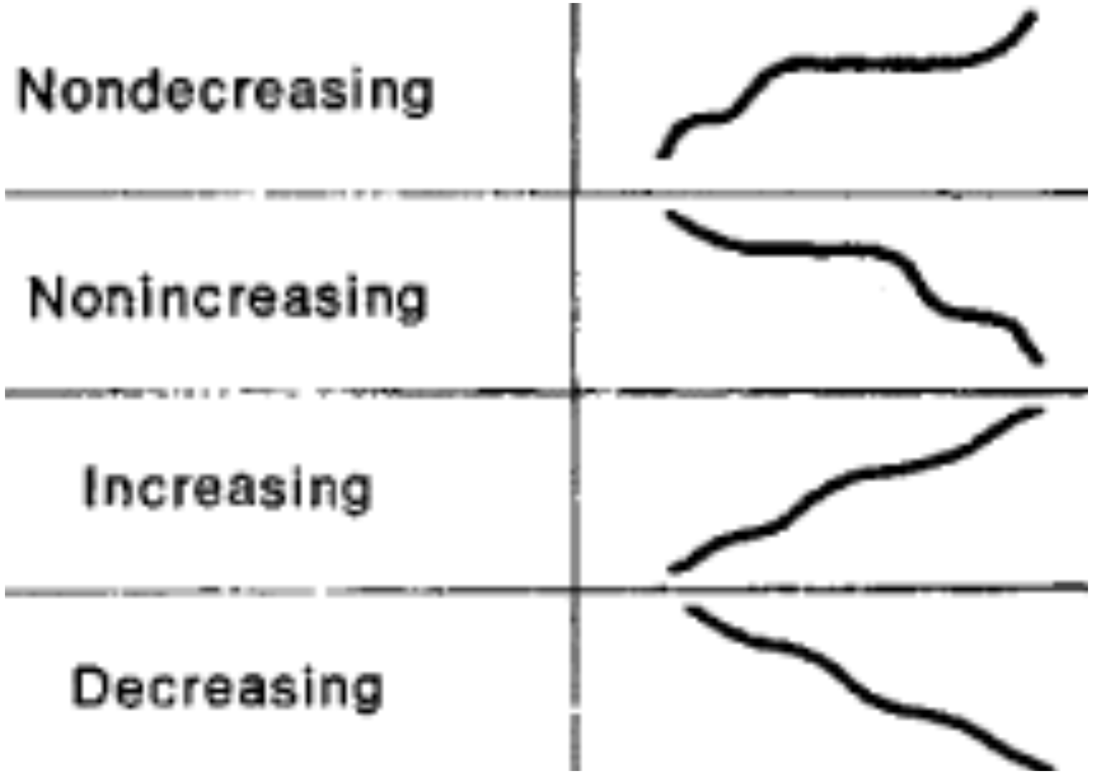
Transformation of Random Variables

Suyash P. Awate

Transformation of Random Variables

- Let X be a (continuous) random variable (RV) with probability density function (PDF) p(X)
- Let continuous function g(.) be strictly monotonically **increasing**
  - If a < b, then g(a) < g(b)
  - We will generalize/extend the class of functions later

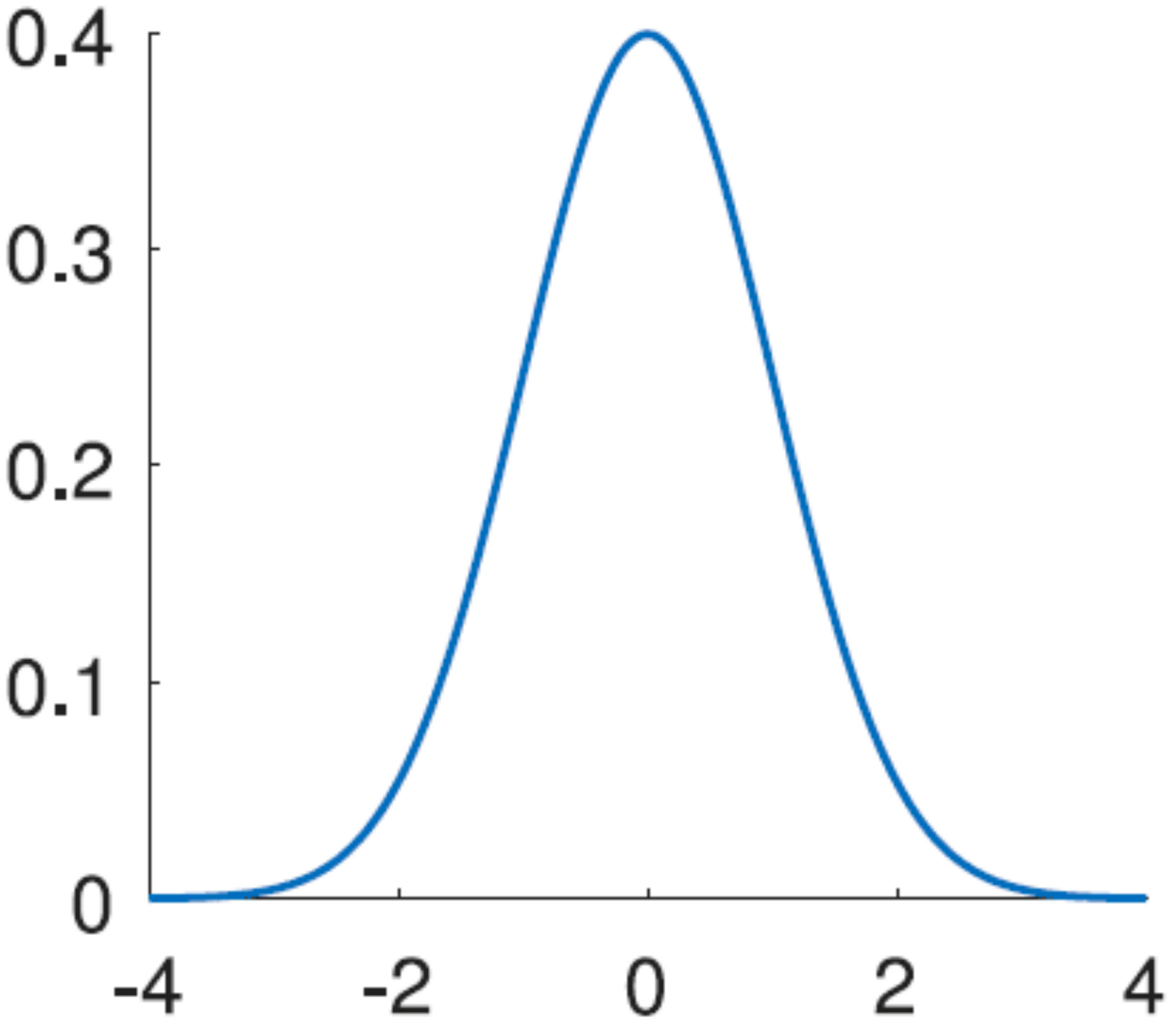


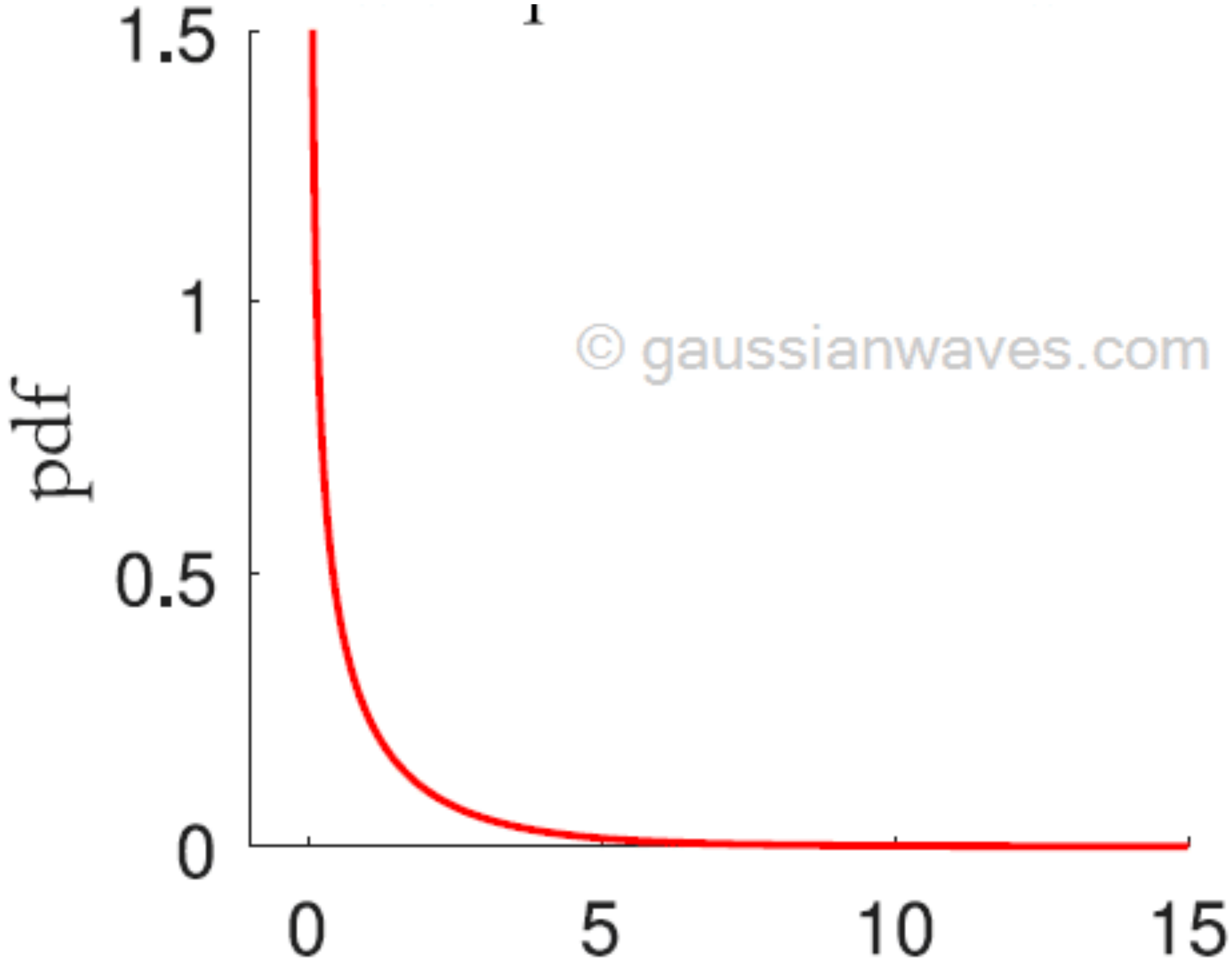


- Consider the **transformed variable** Y := g(X)
- What is the PDF q(Y) of RV Y ?

Transformation of Random Variables

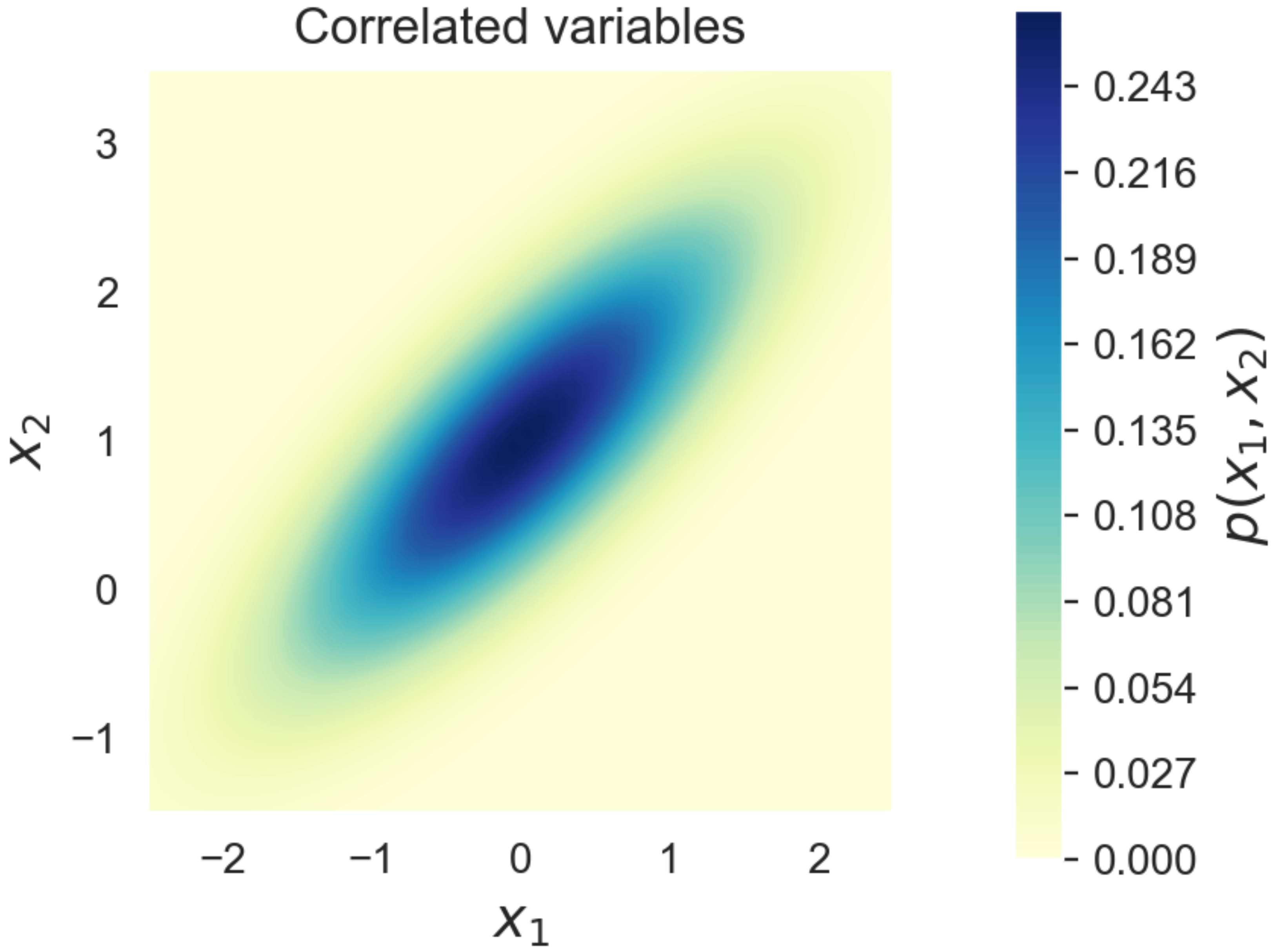
- Example
  - If X has a Normal PDF, then what is the PDF for $Y := X^2$?





Transformation of Random Variables

- Example
  - If
    - RVs U,V are independent Gaussian
    - X1 := aU + bV
    - X2 := cU + dV
  - Then
    - What is p(X1,X2) ?



Transformation of Random Variables

- Principle of probability-mass conservation
  - Consider the events $\{x : x \in (a, b)\}$ and $\{y : y \in (g(a), g(b))\}$
  - Because we assumed that g(.) was increasing, $P(g(a) < Y < g(b)) = P(a < X < b)$

  - So, the probability mass of X in interval (a, b) gets mapped to the probability mass of Y in interval (g(a), g(b))

Now, $P(g(a) < Y < g(b)) := \int_{g(a)}^{g(b)} q(y)dy$

Also, $P(a < X < b) := \int_a^b p(x)dx$

  - Write the second integral in terms of y, using the known relationship y = g(x)

Transformation of Random Variables

- We found that these probabilities → are equal

$$P(g(a) < Y < g(b)) := \int_{g(a)}^{g(b)} q(y)dy$$

$$P(a < X < b) := \int_a^b p(x)dx$$

- We have, $x = g^{-1}(y)$

$$dx = \left(\frac{d}{dy}g^{-1}(y)\right) dy$$

Then, $P(a < X < b) = \int_{g(a)}^{g(b)} p(g^{-1}(y)) \left(\frac{d}{dy}g^{-1}(y)\right) dy$

- This mass conservation holds for **every interval** (a,b), small & large

Thus, $q(y) = p(g^{-1}(y))\frac{d}{dy}g^{-1}(y)$, for all y

  - Note: $g^{-1}(.)$ may be non-differentiable at a countably-many isolated points (and that is okay)

Transformation of Random Variables

- Another similar way to look at it
  - Also using mass conversation, equating integrals
- Equality of CDF values: $P_Y(-\infty \leq Y \leq y) = P_X(-\infty \leq X \leq g^{-1}(y))$
  - **For all y**
- PDF is derivative of CDF
- So, $q(y) = \frac{d}{dy} P_Y(-\infty \leq Y \leq y)$

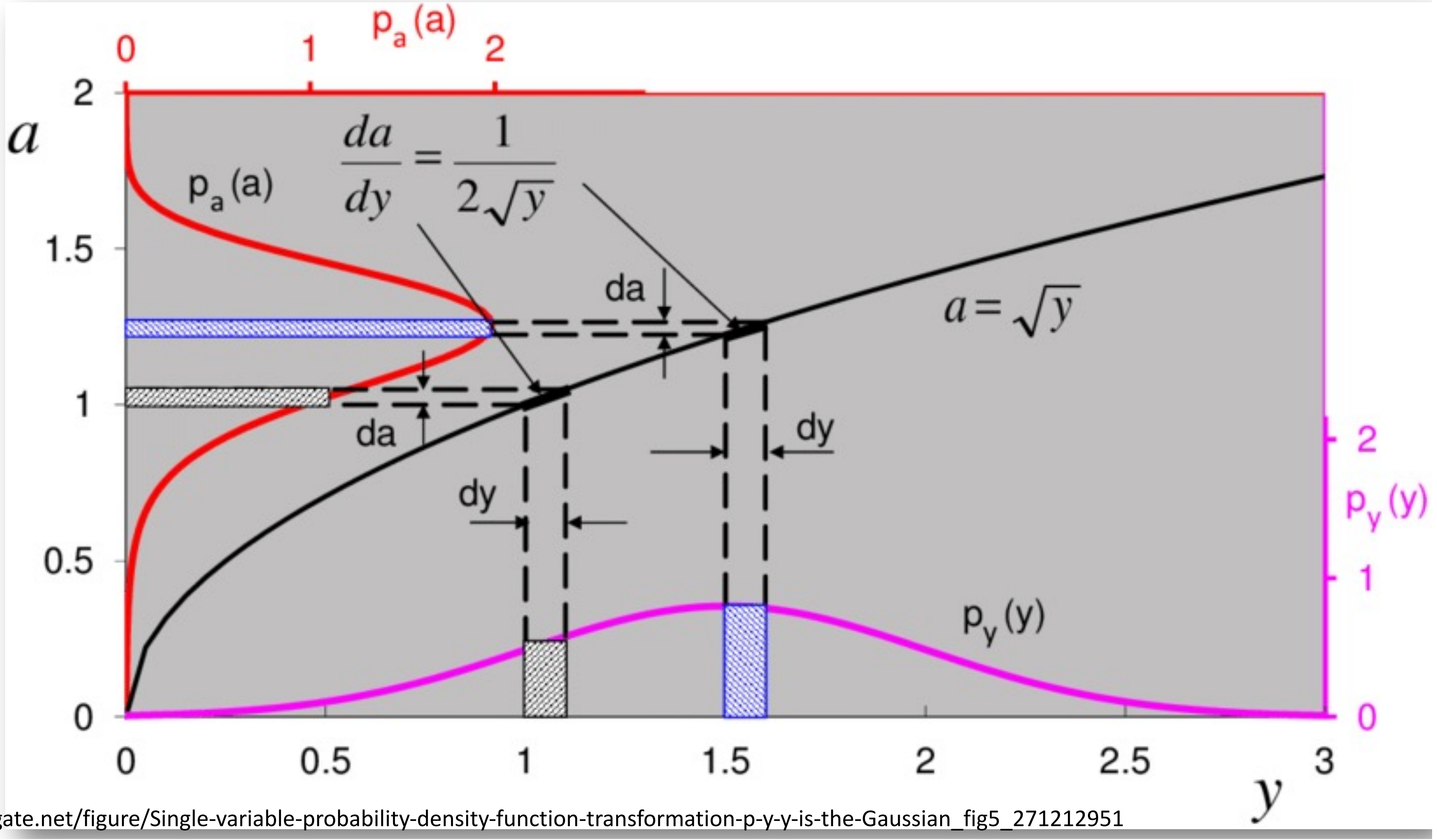
$= \frac{d}{dy} P_X\left(-\infty \leq X \leq g^{-1}(y)\right)$

$= \frac{d}{dy} \int_{-\infty}^{g^{-1}(y)} p(x)dx$

$= p(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$ (using Leibniz integral rule)

Transformation of Random Variables

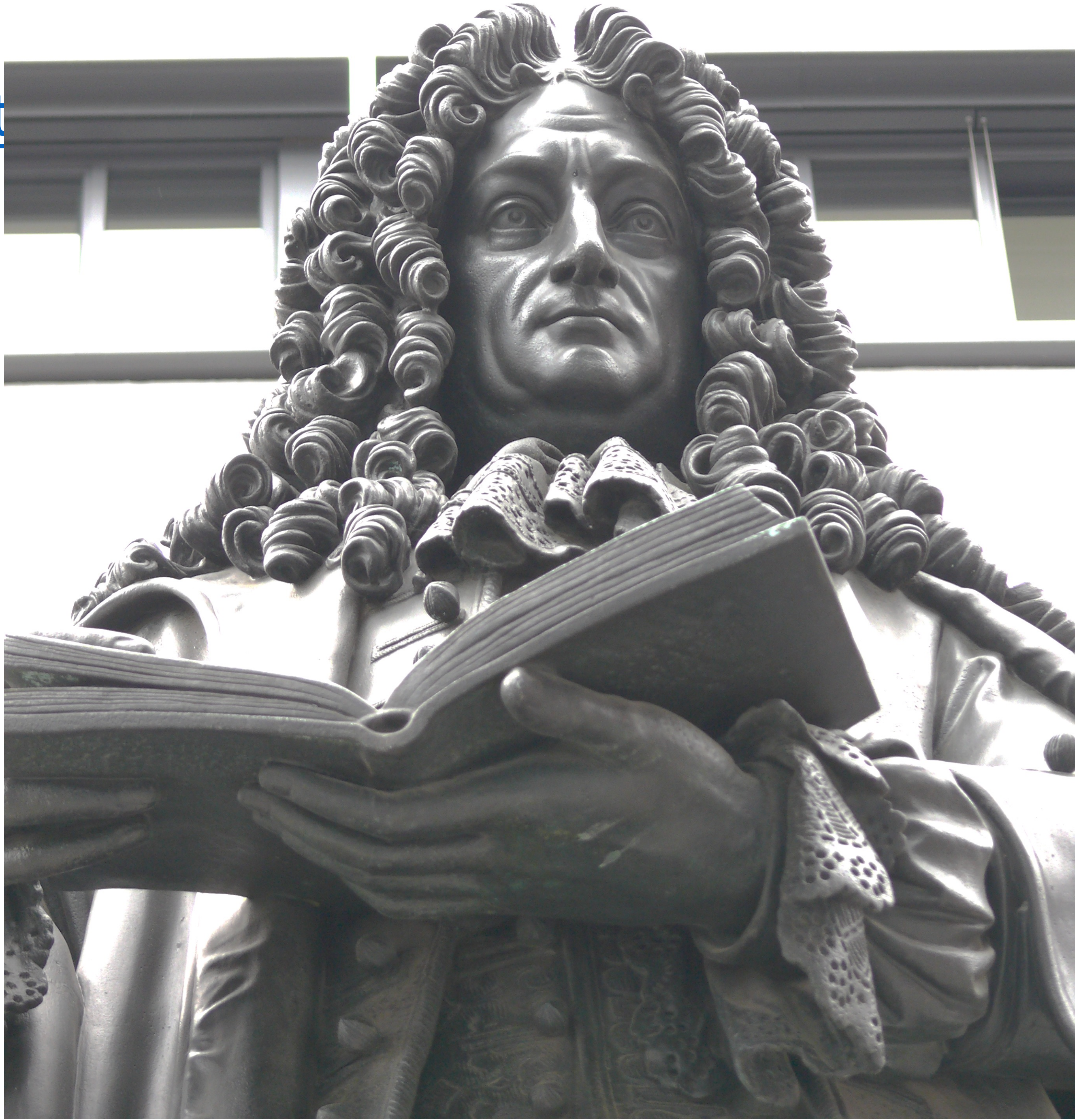
- Example
  - P(Y)
  - A := sqrt(Y)
  - g(.) is sqrt(.)
  - To find P(A)



https://www.researchgate.net/figure/Single-variable-probability-density-function-transformation-p-y-y-is-the-Gaussian_fig5_271212951

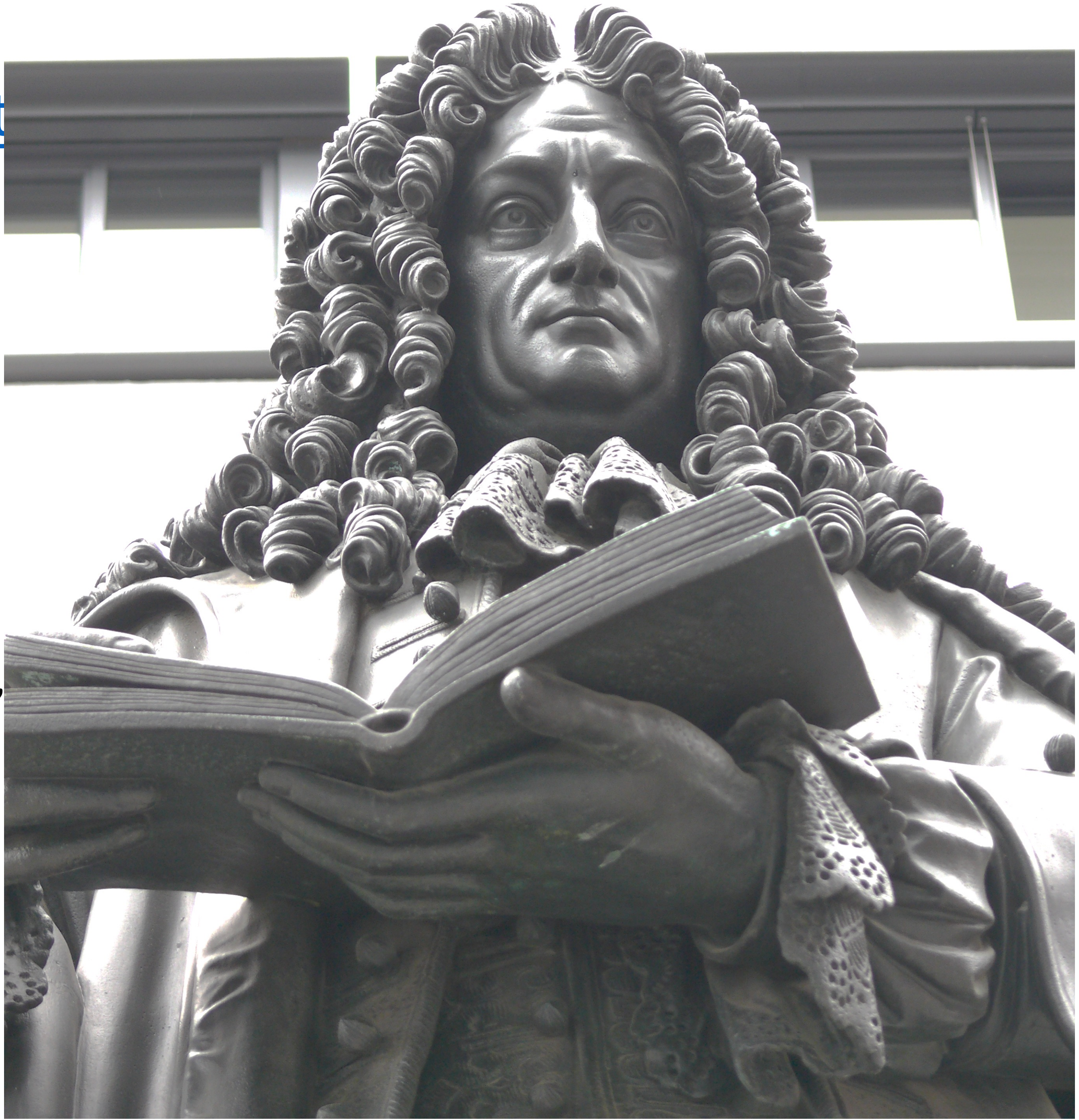
Leibniz

- https://en.wikipedia.org/wiki/Got
  - German polymath
  - Mathematician, philosopher, scientist, diplomat
  - Major contributions to physics, and technology
  - "Anticipated notions that surfaced much later in probability theory, biology, medicine, geology, psychology, linguistics and computer science."



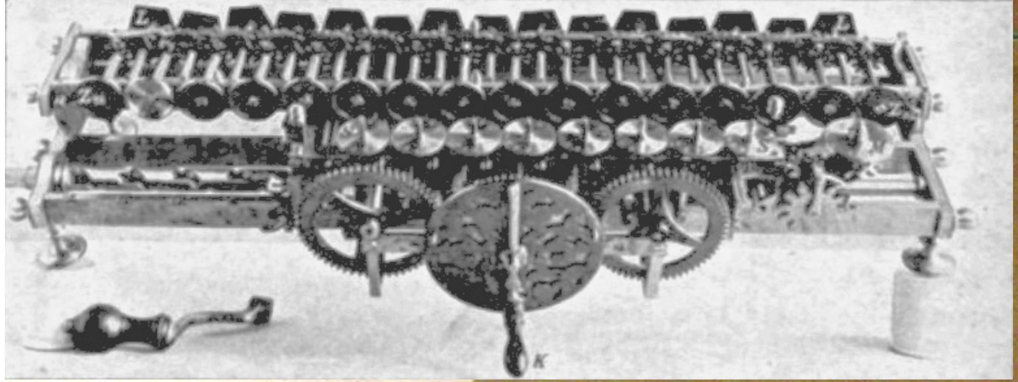
Leibniz

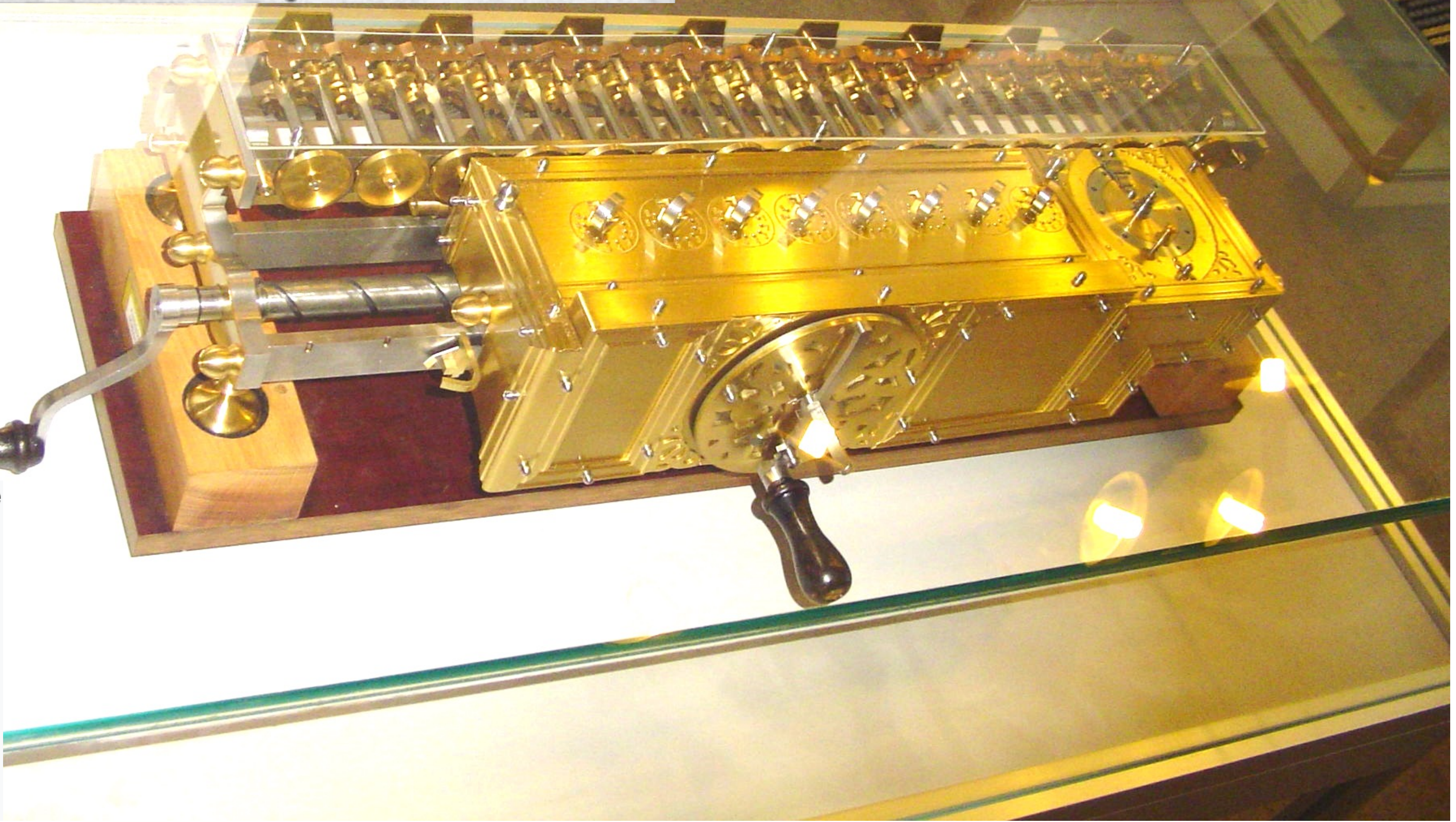
- https://en.wikipedia.org/wiki/Got
  - Developed main ideas of differential and integral calculus, independently of Newton
  - “Founder of computer science”
    - Refined the binary number system
    - Formal logic
  - “In philosophy and theology,
    … most noted for his optimism, i.e.,
    his conclusion that our world is,
    in a qualified sense,
    the best possible world
    that God could have created”



Leibniz

- [Leibniz calculator]
  - Mechanical calculator
  - Invented over 1672-94





Replica of Leibniz's stepped reckoner in the Deutsches Museum.

> ... it is beneath the dignity of excellent men to waste their time in calculation when any peasant could do the work just as accurately with the aid of a machine.
>
> — Gottfried Leibniz[1]

en.wikipedia.org/wiki/List_of_pioneers_in_computer_science

Achievement date	Person	Achievement
500 BC ~	Pāṇini	*Ashtadhyayi* Indian Sanskrit grammar was systematised and technical, using metarules, transformations, and recursions, a forerunner to formal language theory an basis for Panini-Backus form used to describe programming languages.
830~	Al-Khwarizmi	The term algorithm is derived from the algorism, the technique of performing arithmeti with Hindu–Arabic numerals popularised by al-Khwarizmi in his book *On the Calculation with Hindu Numerals*.[1][2][3]
850~	Banū Mūsā	Three brothers who wrote the *Book of Ingenious Devices*, describing what appears to be the first programmable machine, an automatic flute player.[5]
1206	Al-Jazari	Invented programmable machines, including programmable humanoid robots,[33] and the castle clock, an astronomical clock considered the first programmable analog computer[34]
1300~	Llull, Ramon	Designed multiple symbolic representations machines, and pioneered notions of symbolic representation and manipulation to produce knowledge—both of which were major influences on Leibniz.
1642	Pascal, Blaise	Invented the mechanical calculator.
1670~	Leibniz, Gottfried	Made advances in symbolic logic, such as the Calculus ratiocinator, that were heavily influential on Gottlob Frege. He anticipated later developments in first-order predicate calculus, which were crucial for the theoretical foundations of computer science.

Transformation of Random Variables

- We found the relationship between PDF q(.) of Y and PDF p(.) of X, in terms of the strictly-increasing transformation function g(.)

$$q(y) = p(g^{-1}(y))\frac{d}{dy}g^{-1}(y), \text{ for all } y$$

- If g(·) is strictly increasing, then:
  - a < b ⇒ g(a) < g(b)
  - Derivative of g-inverse(.) is positive
  - So, the above formula holds good
- If g(·) is strictly decreasing, then:
  - a < b ⇒ g(a) > g(b)
  - Derivative of g-inverse(.) is negative.
  - What to do then ?

Transformation of Random Variables

- Take the integral limits to go from a smaller number to a larger number

We have, $x = g^{-1}(y)$

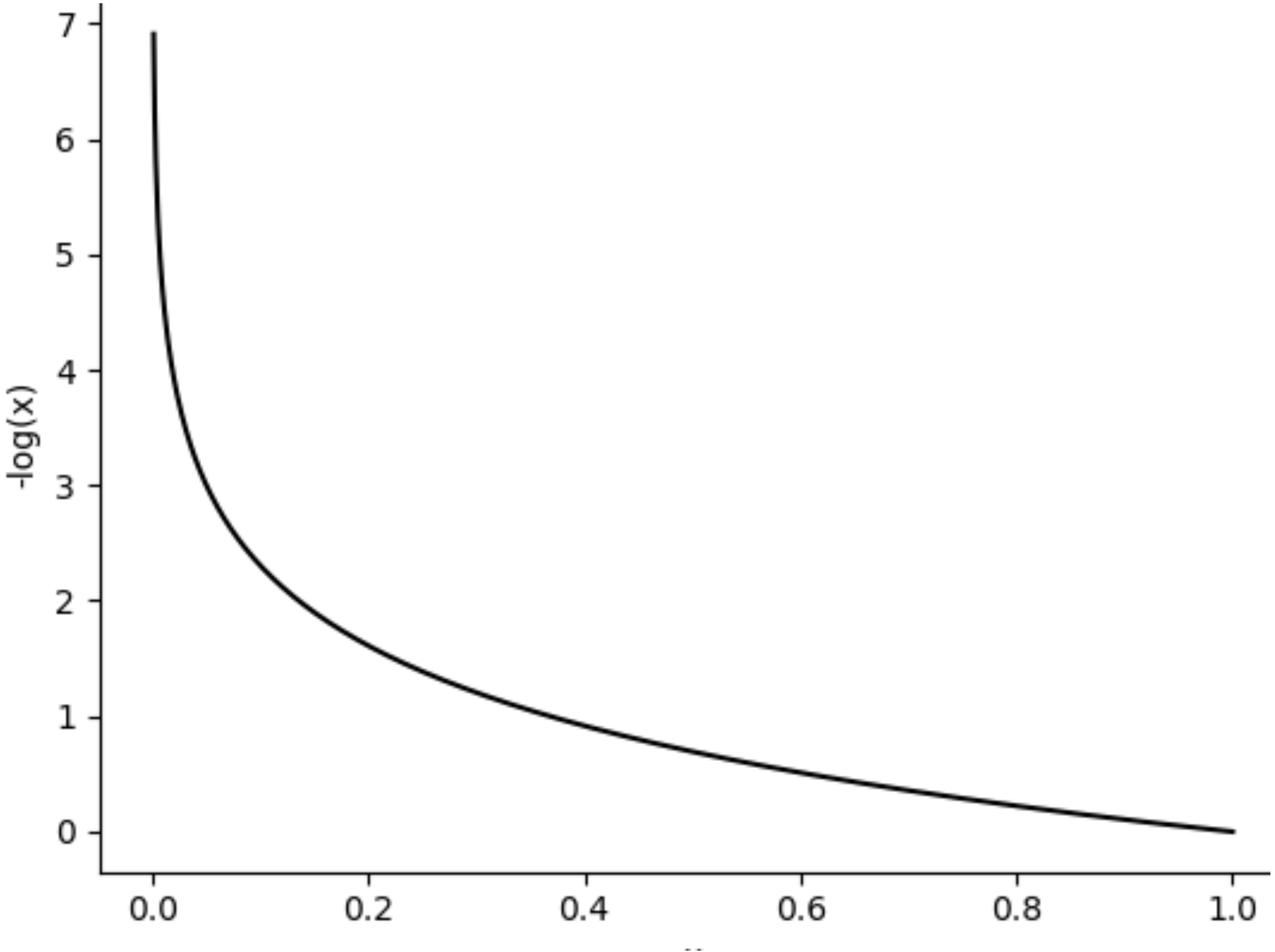
$$dx = \left(\frac{d}{dy} g^{-1}(y)\right) dy$$

Then, $P(a < X < b) = \int_{g(a)}^{g(b)} p(g^{-1}(y)) \left(\frac{d}{dy} g^{-1}(y)\right) dy$

- Notice that mass-conversation principle remains same in tiny intervals, irrespective of whether slope is positive or negative
- For convenience, to handle both cases (increasing/decreasing), we:
  - Write $$q(y) = p(g^{-1}(y)) \left|\frac{d}{dy} g^{-1}(y)\right|$$
  - Taking the absolute value ensures that PDF q(.) is always non-negative

Transformation of Random Variables

- Consider a RV $X \sim U(0,1)$ (generated by the C/C++ rand() function)
- Consider the transformation $Y := (-1/\lambda)\log(X)$, where $\lambda > 0$
- What is $q(Y)$?



Transformation of Random Variables

- Consider a RV X ~ U(0,1) (generated by the C/C++ rand() function)
- Consider the transformation Y := (−1/λ) log(X), where λ > 0
- What is q(Y) ?

$y = -(1/\lambda)\log(x) \implies x = \exp(-\lambda y)$. This is the $g^{-1}(\cdot)$ function.

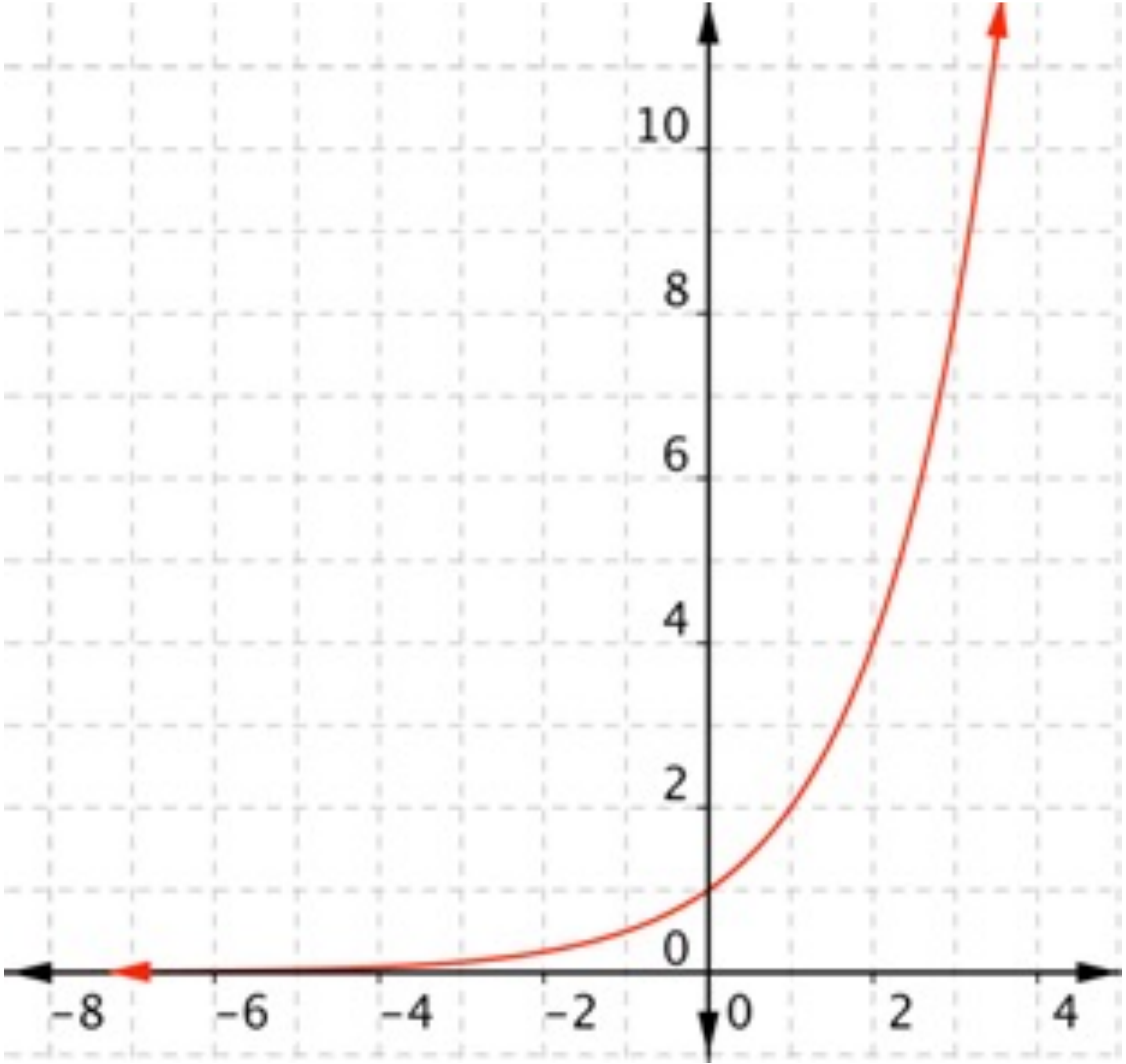
$$\left|\frac{d}{dy}g^{-1}(y)\right| = \lambda\exp(-\lambda y)$$

So, $q(y) = p(g^{-1}(y))\left|\frac{d}{dy}g^{-1}(y)\right| = \lambda\exp(-\lambda y)$

- Thus, Y has the exponential PDF with parameter λ (with mean = 1/λ)

Transformation of Random Variables

- Consider a RV $X \sim U(-a/2, +a/2)$
- Consider $Y := \exp(X)$
- What is $q(Y)$?



Transformation of Random Variables

- Consider a RV X ~ U(−a/2, +a/2)
- Consider Y := exp(X)
- What is q(Y) ?

$y = \exp(x) \implies x = \log(y)$. This is the $g^{-1}(\cdot)$ function.

$$\left|\frac{d}{dy} g^{-1}(y)\right| = 1/y$$

So, $q(y) = p(g^{-1}(y)) \left|\frac{d}{dy} g^{-1}(y)\right| = (1/a)(1/y)$

- Thus, Y has PDF q(y) = 1/(ay) for y ∈ (exp(−a/2), exp(a/2))

Transformation of Random Variables

- Consider a RV X ~ G(0,1) (standard Normal PDF)
- Consider Y := aX, with ‘a’ non-zero
- What is q(Y) ?

$$y := ax \implies x = y/a \implies g^{-1}(y) = y/a$$

$$\left|\frac{d}{dy} g^{-1}(y)\right| = 1/a$$

$$q(y) := p(g^{-1}(y)) \left|\frac{d}{dy} g^{-1}(y)\right| = p\left(\frac{y}{a}\right)\frac{1}{a} = \frac{1}{a\sqrt{2\pi}} \exp\left(-\frac{y^2}{2a^2}\right)$$

- Thus, p(Y) is also a Gaussian with variance σ^2 scaled by a factor of a^2

Transformation of Random Variables

- Consider a RV X $\sim$ G(0,a^2)
- Consider Y := X + b
- What is q(Y) ?

$$y := b + x \implies x = y - b \implies g^{-1}(y) = y - b$$

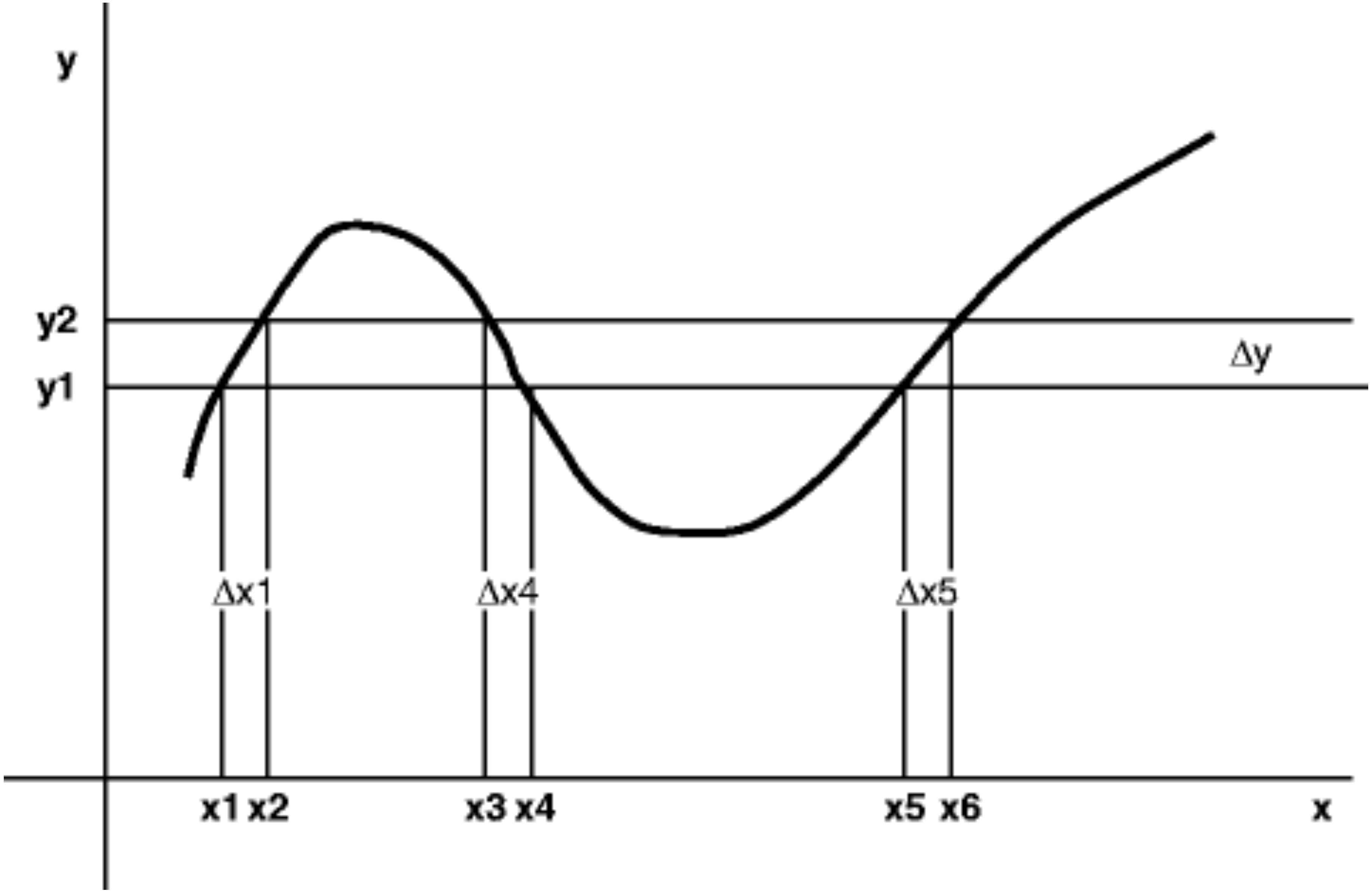
$$\left|\frac{d}{dy} g^{-1}(y)\right| = 1$$

$$q(y) := p(g^{-1}(y)) \left|\frac{d}{dy} g^{-1}(y)\right| = p(y-b) \cdot 1 = \frac{1}{a\sqrt{2\pi}} \exp\left(-\frac{(y-b)^2}{2a^2}\right)$$

- Thus, p(Y) is also a Gaussian with μ translated by b

Transformation of Random Variables

- General non-monotonic functions

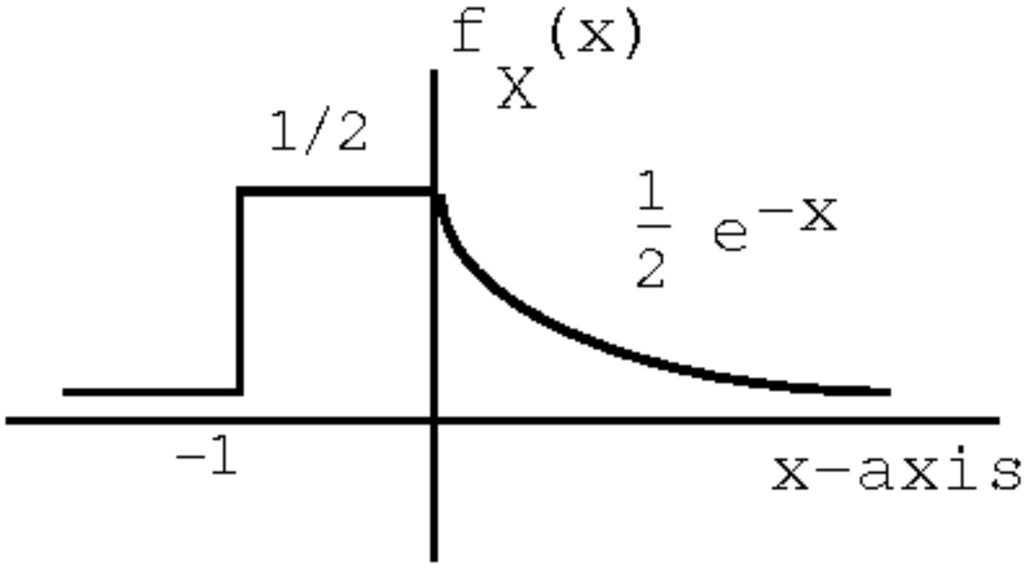


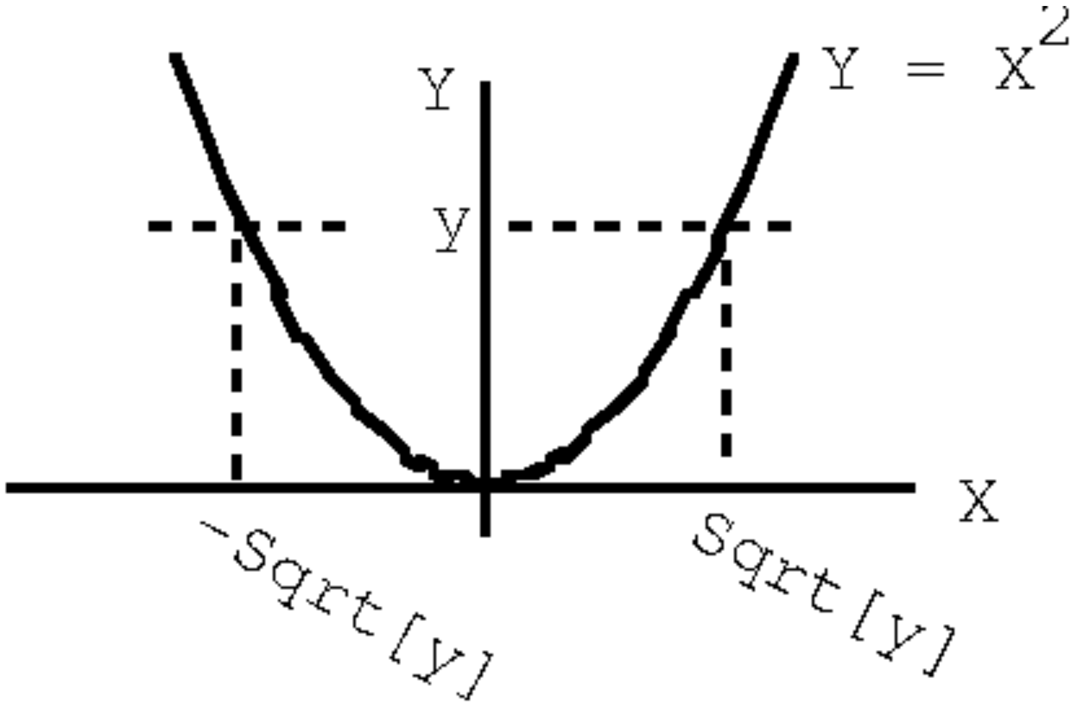
Transformation of Random Variables

- Consider a PDF P(X) as follows:

$$P(x) := 0 \text{ for } x \le -1$$
$$P(x) := 0.5 \text{ for } x \in (-1, 0)$$
$$P(x) := 0.5 \exp(-x) \text{ for } x \ge 0$$

- Consider a transformation function Y := g(X) := X^2
- What is PDF q(y) of Y ?





Transformation of Random Variables

- Consider a PDF P(X) as follows:

$$P(x) := 0 \text{ for } x \leq -1$$
$$P(x) := 0.5 \text{ for } x \in (-1, 0)$$
$$P(x) := 0.5 \exp(-x) \text{ for } x \geq 0$$

- Consider a transformation function Y := g(X) := X^2
- What is PDF q(y) of Y ?

$$y := x^2 \implies x = \pm\sqrt{y} \implies g^{-1}(y) = \pm\sqrt{y}$$

$$\left|\frac{d}{dy} g^{-1}(y)\right| = \frac{1}{2\sqrt{y}}$$

Transformation of Random Variables

- Consider a PDF P(X) as follows:

$$P(x) := 0 \text{ for } x \leq -1$$
$$P(x) := 0.5 \text{ for } x \in (-1, 0)$$
$$P(x) := 0.5 \exp(-x) \text{ for } x \geq 0$$

- Consider a transformation function Y := g(X) := X^2
- What is PDF q(y) of Y ?

Case 1: $x \in (-1, 0)$. In this case, $g(\cdot)$ is a *decreasing* function. Mass conservation applies.

$$\text{For } y \in (0, 1) : q_1(y) := p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = (0.5) \frac{1}{2\sqrt{y}} = \frac{1}{4\sqrt{y}}$$

Case 2: $x \geq 0$. In this case, $g(\cdot)$ is a *increasing* function. Mass conservation applies.

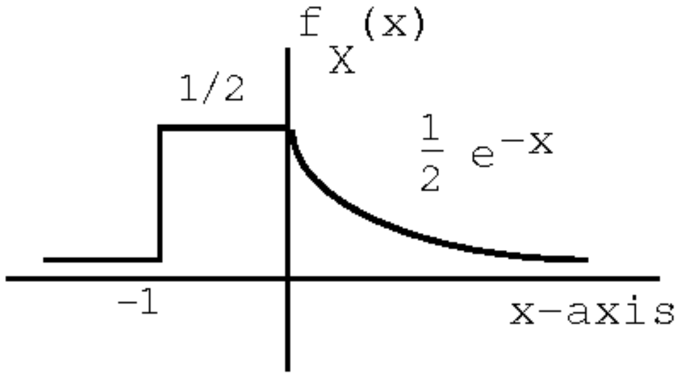
$$\text{For } y \geq 0 : q_2(y) := p(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| = (0.5 \exp(-\sqrt{y})) \frac{1}{2\sqrt{y}} = \frac{\exp(-\sqrt{y})}{4\sqrt{y}}$$

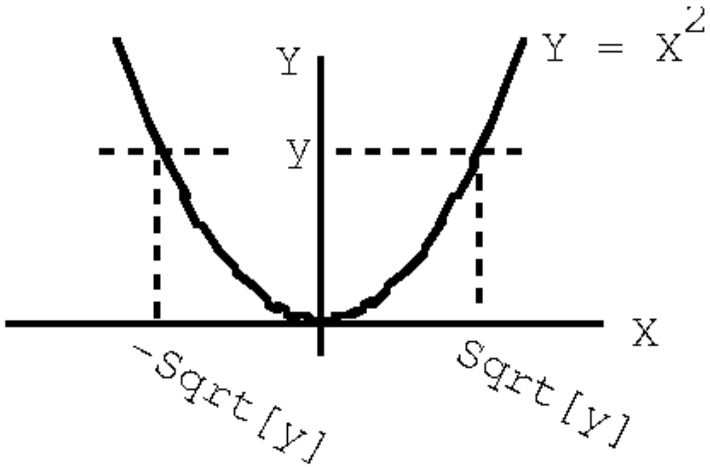
Transformation of Random Variables

- Consider a PDF P(X) as follows:

$$P(x) := 0 \text{ for } x \le -1$$
$$P(x) := 0.5 \text{ for } x \in (-1, 0)$$
$$P(x) := 0.5 \exp(-x) \text{ for } x \ge 0$$

- Consider a transformation function Y := g(X) := X^2
- What is PDF q(y) of Y ?
- Desired PDF q(y) = $q_1(y) + q_2(y)$
- In the region y ∈ (0,1), probability mass comes from Case 1 & Case 2





Transformation of Random Variables

- Consider a PDF P(X) as follows:

$$P(x) := 0 \text{ for } x \leq -1$$
$$P(x) := 0.5 \text{ for } x \in (-1, 0)$$
$$P(x) := 0.5 \exp(-x) \text{ for } x \geq 0$$

- Consider a transformation function Y := g(X) := X^2
- What is PDF q(y) of Y ?

Thus,

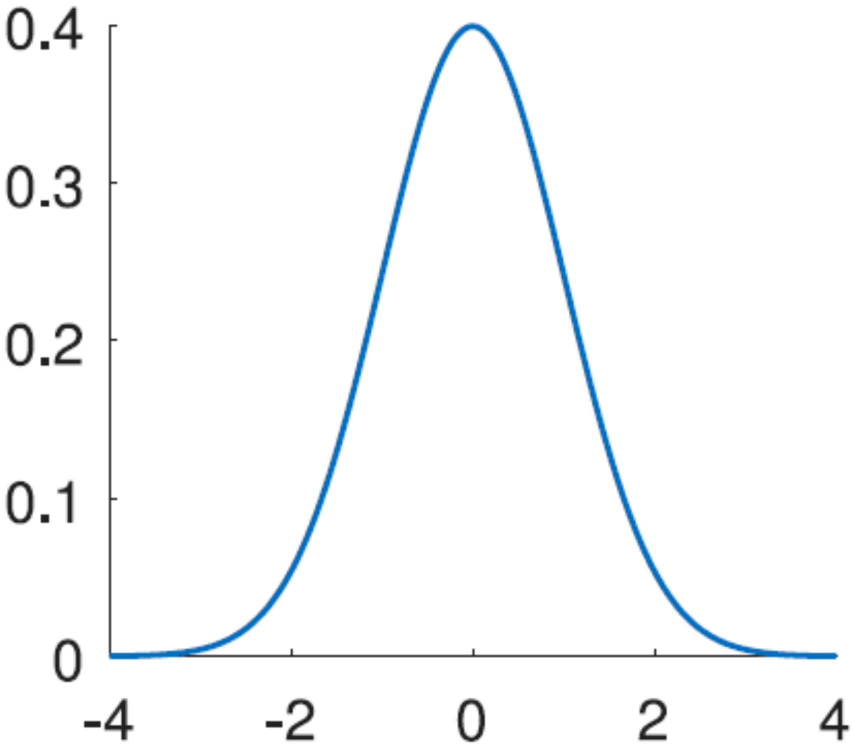
(i) for $y \in (0,1)$, PDF $q(y) = \frac{1}{4\sqrt{y}}(1 + \exp(-\sqrt{y}))$

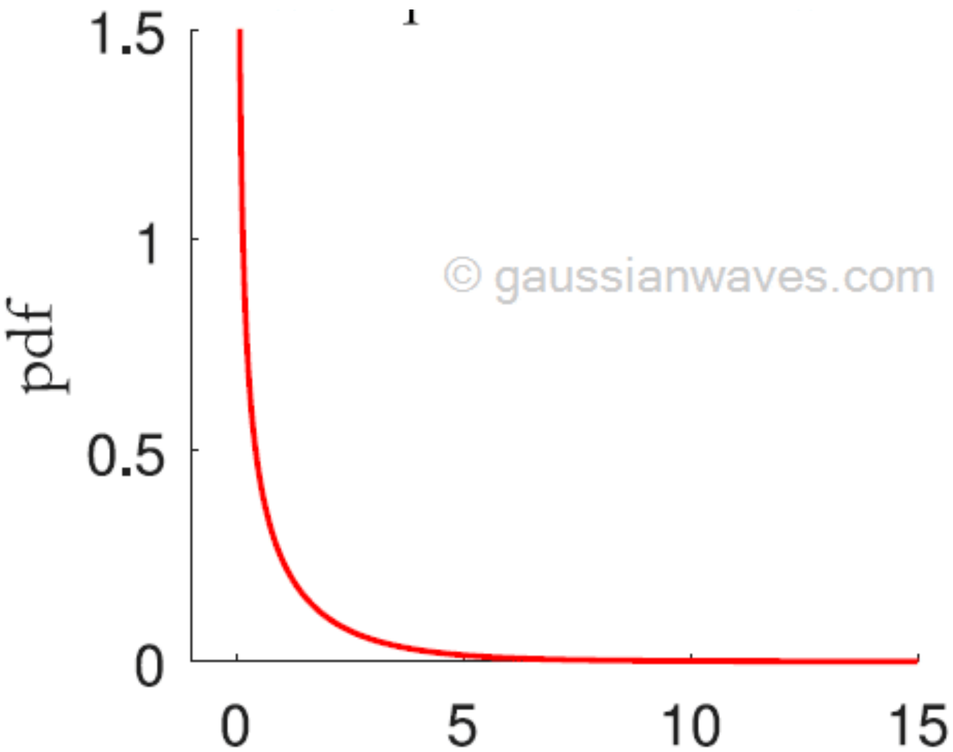
(ii) for $y \geq 1$, PDF $q(y) = \frac{\exp(-\sqrt{y})}{4\sqrt{y}}$

- There will be a jump discontinuity at y = 1,
  where left limit = (1+exp(-1))/4 and right limit = exp(-1)/4

Transformation of Random Variables

- Let $X \sim G(0,1)$
- Let $Y := X^2$
- What is $P(Y)$, defined as the chi-square PDF ?





Transformation of Random Variables

- Let X ~ G(0,1)
- Let Y := X^2
- What is P(Y) ?

$$y := x^2 \implies x = \pm\sqrt{y} \implies g^{-1}(y) = \pm\sqrt{y}$$

$$\left|\frac{d}{dy}g^{-1}(y)\right| = \frac{1}{2\sqrt{y}}$$

- Case 1: x ≤ 0. Here, g(.) is a decreasing function

For $y \geq 0 : q_1(y) := p(g^{-1}(y))\left|\frac{d}{dy}g^{-1}(y)\right| = \frac{\exp(-0.5(\sqrt{y})^2)}{\sqrt{2\pi}}\frac{1}{2\sqrt{y}} = \frac{\exp(-0.5y)}{2\sqrt{y2\pi}}$

- Case 2: x > 0. Here, g(·) is a increasing function

For $y > 0 : q_2(y) := \frac{\exp(-0.5y)}{2\sqrt{y2\pi}}$

- Desired chi-square PDF: $q(y) = q_1(y) + q_2(y) = (1/\mathrm{sqrt}(y2\pi))\exp(-0.5y)$

Transformation of Random Variables

- Let X have a Gamma PDF,

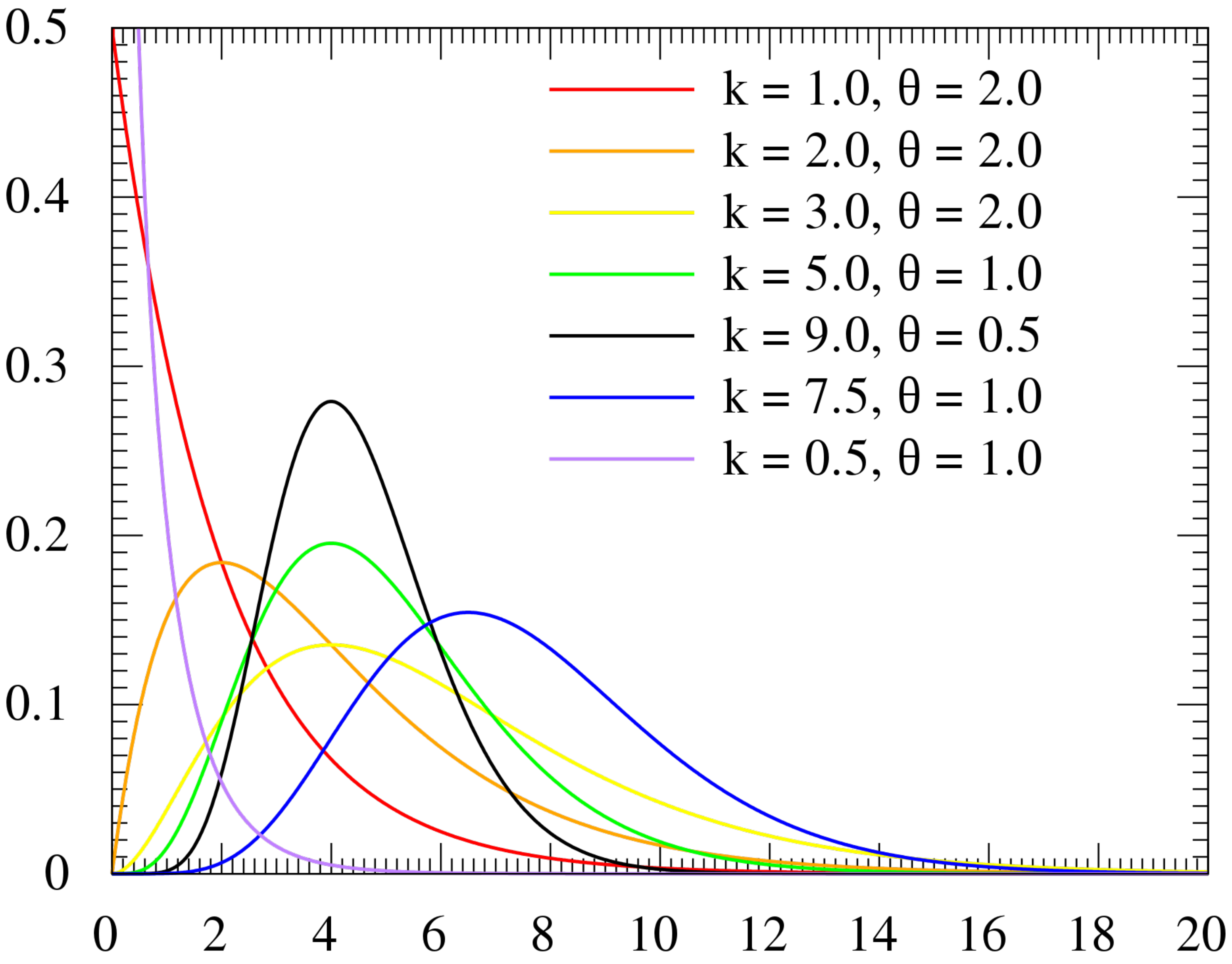
$$P(x) = \mathsf{Gamma}(x|\alpha,\beta) = (\beta^\alpha/\Gamma(\alpha))x^{\alpha-1}\exp(-\beta x)$$

where α (shape) > 0, β (rate) > 0, x > 0, Γ(·) = gamma function defined for all complex numbers with real part positive

  - $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x}\, dx, \qquad \Re(z) > 0$
  - For positive integer n, gamma(n) = factorial(n-1)

Transformation of Random Variables

- Gamma PDF
  - k = shape = α
  - theta = scale = 1/rate = 1/ β
  - [Link]



Transformation of Random Variables

- Let X have a Gamma PDF,

$$P(x) = \text{Gamma}(x|\alpha, \beta) = (\beta^\alpha / \Gamma(\alpha)) x^{\alpha-1} \exp(-\beta x)$$

where α (shape) > 0, β (rate) > 0, x > 0, Γ(·) = gamma function defined for all complex numbers with real part positive

  - $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x}\, dx, \qquad \Re(z) > 0$
  - For positive integer n, gamma(n) = factorial(n-1)
- Consider the transformation Y := 1/X
- What is the PDF of Y ?

$$y := 1/x \implies x = 1/y \implies g^{-1}(y) = 1/y$$

$$\left| \frac{d}{dy} g^{-1}(y) \right| = \frac{1}{y^2} \text{ for } y > 0$$

Transformation of Random Variables

- Let X have a Gamma PDF,

$$P(x) = \text{Gamma}(x|\alpha, \beta) = (\beta^\alpha/\Gamma(\alpha))x^{\alpha-1}\exp(-\beta x)$$

where α > 0, β > 0, x > 0, and Γ(·) is the well-known gamma function defined for all complex numbers with real part positive.

- Consider the transformation Y := 1/X
- What is the PDF of Y ?

$$y := 1/x \implies x = 1/y \implies g^{-1}(y) = 1/y$$

$$\left|\frac{d}{dy}g^{-1}(y)\right| = \frac{1}{y^2} \text{ for } y > 0$$

$$q_1(y) := p(g^{-1}(y))\left|\frac{d}{dy}g^{-1}(y)\right| = (\beta^\alpha/\Gamma(\alpha))y^{1-\alpha}\exp(-\beta/y)\frac{1}{y^2} = (\beta^\alpha/\Gamma(\alpha))y^{-\alpha-1}\exp(-\beta/y)$$

  - This is called the inverse-Gamma PDF

Transformation of Random Variables

- Inverse-Gamma PDF