Probability & Statistics

September 7

$\mathbb{R}$, $(\Omega, \mathcal{F}, P)$, — discrete & continuous r.v.s. $\longleftrightarrow$ pmf $\longleftrightarrow$ CDF

$\longleftrightarrow$ pdf $\longleftrightarrow$ CDF

[Expectation, median, variance]

Suppose we have a r.v. X.

Let $f_X(x)$ be the pdf/pmf of X.

Let $Y = X^2$ (transformation of r.v.)

Ex: $X \sim$ uniform (10) $R_X = \{0, 1, 2, \ldots, 9\}$

$X=x$	0	1	2	...	9
$f(x)$	$1/10$	$1/10$	$1/10$		$1/10$

← pmf of X

$Y=y$	0	1	4	9	....	81
$f(y)$	$1/10$	$1/10$	$1/10$			$1/10$

← pmf of $Y = X^2$

Ex: $X \sim \text{Uniform}(11)$ $R_X = \{-5, -4, \dots, -2, -1, 0, 1, 2, \dots, 5\}$

pmf of X =

$X = x$	-5	-4	-3		0	1		5
$f(x)$	$1/11$	$1/11$	$1/11$					$1/11$

pmf of Y =

$Y = y = x^2$	0	1	4	9	16	25
$f(y)$	$1/11$	$2/11$	$2/11$	$2/11$	$2/11$	$2/11$

Let X be a continuous r.v. with pdf $f_X(x)$.

Find the density of r.v. $Y = X^2$.

Let $F_X(x)$ be the CDF of X and let

$G_Y(y)$ be the CDF of Y.

For any real number $y \in \mathbb{R}$

$$P(a \le X \le b) = F_X(b) - F_X(a), \quad a = -\sqrt{y}, \quad b = \sqrt{y}$$

$$G_Y(y) = P(Y \le y)$$

$$= P(X^2 \le y)$$

$$= P(-\sqrt{y} \le X \le \sqrt{y})$$

$$G_Y(y) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

Let $g_Y(y)$ denote pdf of Y.

$$g_Y(y) = \frac{d}{dy} G_Y(y)$$

$$g_Y(y) = \frac{d}{dy} G_Y(y)$$

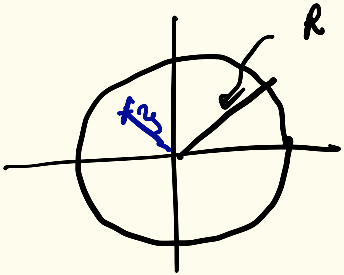
$$= \frac{d}{dy} \left[F_X(\sqrt{y}) - F_X(-\sqrt{y}) \right)$$

$$= \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y})$$

$$\boxed{g_Y(y) = \frac{1}{2\sqrt{y}} \left[f_X(\sqrt{y}) + f_X(-\sqrt{y}) \right] \qquad \text{for } y \geq 0}$$

Ex:

$$f_X(x) = \begin{cases} \dfrac{2x}{R^2} & 0 \le x \le R \\ 0 & \text{otherwise} \end{cases}$$



Find the density of $Y = X^2$

Let $g_Y(y)$ be the density of Y.

$$g_Y(y) = \frac{1}{2\sqrt{y}} \left[f_X(\sqrt{y}) + f_X(-\sqrt{y}) \right] \qquad 0 \le y \le R^2$$

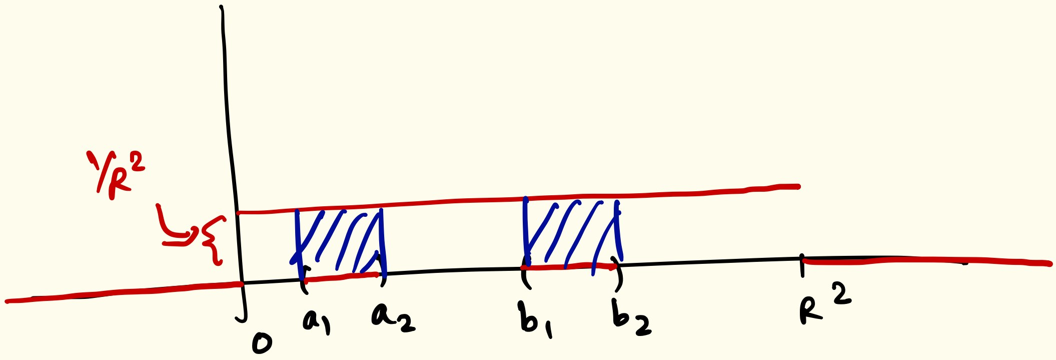
$$0 \le \sqrt{y} \le R$$

$$= \frac{1}{2\sqrt{y}} \left[\frac{2\sqrt{y}}{R^2} + 0 \right]$$

$$\boxed{g_Y(y) = \frac{1}{R^2} \quad 0 \le y \le R^2; \qquad = 0 \quad \text{o.w.}}$$

$$g_Y(y) = \begin{cases} \dfrac{1}{R^2} & 0 \le y \le R^2 \\ 0 & \text{otherwise} \end{cases}$$

← uniform density (continuous)



If one chooses two subsets of equal "length" (size) of $[0, R^2]$, then their probabilities will be same.

$$P(a_1 \le X \le a_2) = P(b_1 \le X \le b_2) \quad \text{given } a_2 - a_1 = b_2 - b_1$$

Continuous uniform density

Let X be a continuous r.v. with pdf defined as

$$f_X(x) = \frac{1}{b-a} \qquad a \le x \le b$$

$$= 0 \qquad \text{otherwise.}$$

for any $a, b \in \mathbb{R}$ s.t. $a < b$.

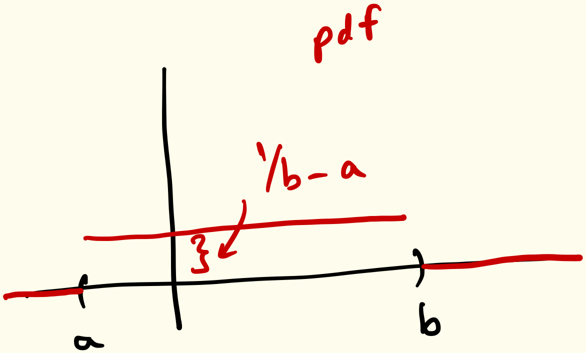
Then X is said to follow uniform density with parameters a & b.

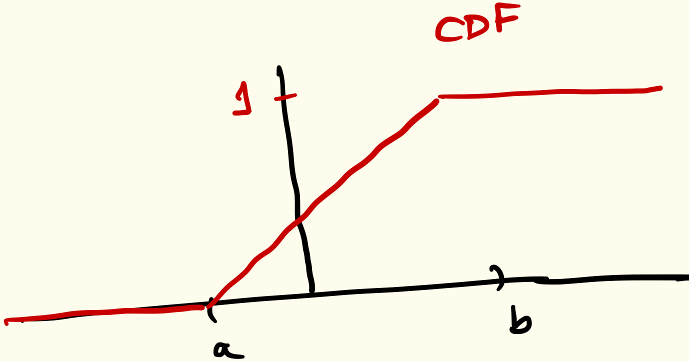
$X \sim$ uniform (a,b) ; $X \sim U(a,b)$

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx = \frac{x-a}{b-a} \qquad a < x < b$$

$$F_X(x) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a \le x \le b \\ 1 & x > b \end{cases}$$

$\boxed{\begin{matrix} a=0 \\ b=1 \end{matrix}}$





Particular case : $a=0,\ b=1$

$X \sim U[0,1]$

$$f_X(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{o.w.} \end{cases}$$

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x \le 1 \\ 1 & x > 1 \end{cases}$$

Ex: $X \sim U[0,1]$

consider the transformation

$$Y = \boxed{-\frac{1}{\lambda}} \log(1-X) \quad \text{for } \lambda > 0$$

Let $G_Y(y)$ be the CDF of Y and $g_Y(y)$ be the pdf of Y.

$$
\begin{aligned}
G_Y(y) &= P(Y \leq y) \\
&= P\left(\boxed{-\frac{1}{\lambda}\log(1-X)} \leq y\right) \quad \text{rearrangement} \\
&= P(\log(1-X) \geq -\lambda y) \quad \text{monotonically } \uparrow \\
&= P(1-X \geq e^{-\lambda y}) \\
G_Y(y) &= P(X \leq 1-e^{-\lambda y}) = F_X(1-e^{-\lambda y})
\end{aligned}
$$

$$G_Y(y) = F_X(1-e^{-\lambda y})$$

$$= 1 - e^{-\lambda y} \quad \forall\ y > 0$$

$$\Rightarrow \quad g_Y(y) = \frac{d}{dy}(1-e^{-\lambda y})$$

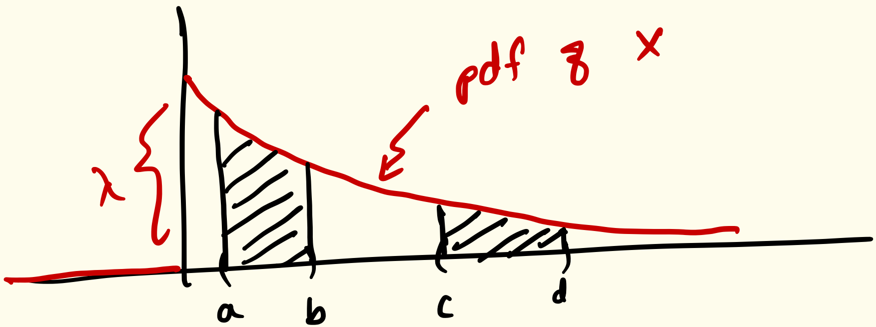
$$= \lambda e^{-\lambda y} \quad \text{for } y > 0$$

$$\boxed{\begin{aligned} g_Y(y) &= \lambda e^{-\lambda y} && \text{for } y>0 \\ &= 0 && \text{otherwise} \end{aligned}}$$

This density is called as exponential density with parameter λ. $(Y \sim \exp(\lambda))$

Ex: Let $X \sim \exp(\lambda)$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \le 0 \end{cases}$$



$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$$

Interesting thing to observe!! (Memoryless property)

For any real numbers $x_0, y_0 > 0$

$$P(X > x_0 + y_0) = 1 - F_X(x_0 + y_0) = 1 - \left(1 - e^{-\lambda(x_0 + y_0)}\right)$$

$$= e^{-\lambda(x_0 + y_0)} = e^{-\lambda x_0} e^{-\lambda y_0}$$

$$= P(X > x_0)\, P(Y > y_0)$$

$$\Rightarrow \frac{P(X > x_0 + y_0)}{P(X > x_0)} = P(X > y_0)$$

$$\Rightarrow \boxed{P(X > x_0 + y_0 \mid X > x_0) = P(X > y_0)}$$

(Ex: Discrete pmf geometric(p) also has memoryless property).

Theorem: Let φ be a differentiable function which is strictly increasing or strictly decreasing on an interval I. Let $\varphi(I)$ denote the range of φ on I and let φ^{-1} be the inverse of φ on I. Let X be a continuous r.v. having density $f_X(x)$ such that $f_X(x) \neq 0$ for $x \in I$. Define $Y = \varphi(X)$. Then the density of Y is given by

$$g(y) = f(\varphi^{-1}(y)) \left| \frac{d}{dy} \varphi^{-1}(y) \right| \qquad y \in \varphi(I)$$

Proof: Let $G(y)$ be the CDF of y.

$$G(y) = P(Y \leq y)$$
$$= P(\varphi(X) \leq y) \qquad [\because \varphi \uparrow$$
$$= P(X \leq \varphi^{-1}(y))$$

$$G_Y(y) = F_X(\varphi^{-1}(y))$$
$$\Rightarrow \boxed{g(y) = f_X(\varphi^{-1}(y)) \frac{d}{dy} \varphi^{-1}(y)}$$

In case of φ to be monotonotically decreasing f?

$$G_Y(y) = P(Y \leq y) = P(\varphi(X) \leq y) = P(X \geq \varphi^{-1}(y))$$
$$G_Y(y) = 1 - F_X(\varphi^{-1}(y)) \Rightarrow \boxed{g_Y(y) = -f_X(\varphi^{-1}(y)) \frac{d}{dy} \varphi^{-1}(y)}$$

Ex: Let X be a continuous r.v. with density $f_X(x)$. Let $a, b \in \mathbb{R}$ and $b \neq 0$.

Define $Y = a + bX = \varphi(X)$ (previous thm)

Compute density of Y.

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

$$\varphi(x) = a + bx \Rightarrow x = \frac{y-a}{b} = \varphi^{-1}(y)$$

$$\boxed{g(y) = f\left(\frac{y-a}{b}\right) \cdot \frac{1}{|b|}}$$

Ex: $f_X(x) = \begin{cases} \frac{2x}{R^2} & 0 \le x \le R \\ 0 & \text{o.w.} \end{cases}$

$b = \frac{1}{R}$
$a = 0$

$0 < y < 1 \leftarrow R_Y$

Consider $Y = \frac{X}{R}$

Here $a = 0$, $b = \frac{1}{R}$

$$g(y) = \frac{1}{|b|} f\left(\frac{y-a}{b}\right)$$

$$= \frac{1}{1/R} \frac{2\left(y/_{1/R}\right)}{R^2}$$

$$\Rightarrow \boxed{\begin{aligned} g(y) &= 2y & 0 \le y \le 1 \\ &= 0 & \text{o.w.} \end{aligned}}$$

Symmetric densities:

f is symmetric density if $f(x) = f(-x)$ $\forall x \in \mathbb{R}$.

Clearly, $U(-a, a)$ is a symmetric density for $a \in \mathbb{R}$.

A random variable X is symmetric if it pdf $f_X(x)$ is symmetric density.

Ex: If X is a symmetric r.v. with CDF $F_X(x)$. Then $F_X(0) = \frac{1}{2}$.

$$F_X(-x) = \int_{-\infty}^{-x} f(y)\,dy = \int_{x}^{\infty} f(-y)\,dy$$

$$= \int_{x}^{\infty} f(y)\,dy \qquad \because f \text{ is symmetric.}$$

$$= 1 - F_X(x)$$

$$\Rightarrow \boxed{F_X(-x) = 1 - F_X(x)} \qquad \forall\, x \in \mathbb{R}$$

For $x = 0$, $F_X(0) = 1 - F_X(0)$

$$\Rightarrow 2F_X(0) = 1 \Rightarrow \boxed{F_X(0) = \frac{1}{2}}$$

Another approach to construct densities.

Ex:

$$g(x) = \frac{1}{1+x^2} \qquad -\infty < x < \infty$$

be a pdf ??

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,dx = \lim_{c\to\infty} \tan^{-1} x \Big|_{-c}^{c} = \pi$$

Cauchy density.

Symmetric.

Then clearly,

$$\boxed{g(x) = \frac{1}{\pi} \frac{1}{1+x^2} \qquad -\infty < x < \infty}$$

is a density.

Ex:

$$f(x) = e^{-x^2/2} \qquad -\infty < x < \infty$$

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{\infty} e^{-x^2/2}\,dx = \sqrt{2\pi} \checkmark$$

$$c = \int_{-\infty}^{\infty} e^{-x^2/2}\,dx$$

$$\Rightarrow c^2 = \int_{-\infty}^{\infty} e^{-x^2/2}\,dx \int_{-\infty}^{\infty} e^{-y^2/2}\,dy$$

$$c^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)/2}\,dx\,dy = \int_0^{\infty}\int_0^{2\pi} r\,e^{-r^2/2}\,d\theta\,dr$$

$$2\pi$$

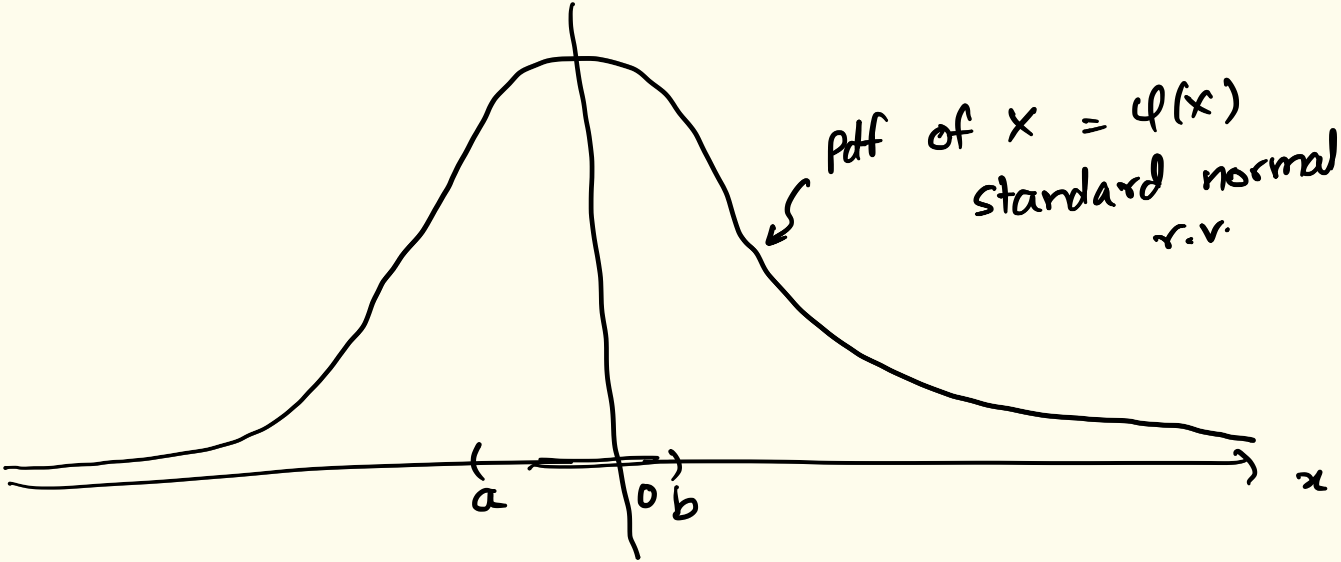
$$\shortparallel$$

Define:

$$\boxed{\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \qquad -\infty < x < \infty}$$

This density is called as standard normal density. The random variable corresponding to this density is called as standard normal random variable.

standard normal density is symmetric.

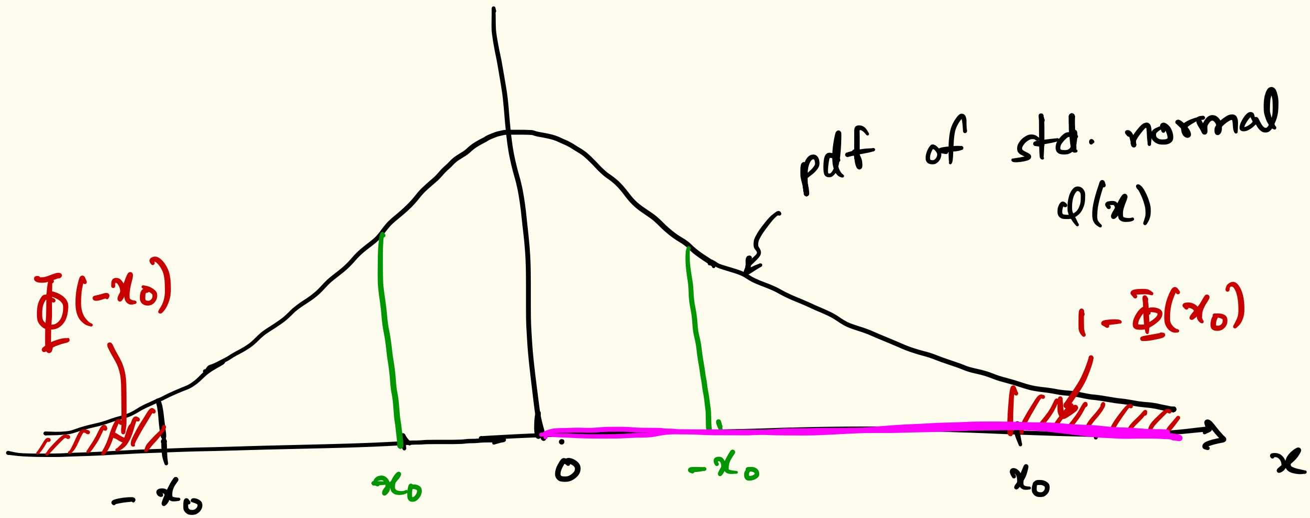


$R_X = \mathbb{R}$ as $f_X(x) > 0 \quad \forall x \in \mathbb{R}$

$$P(a \le X \le b) = F_X(b) - F_X(a)$$

Let $\Phi_X(x)$ denote the CDF of X

$$P(a \le X \le b) = \Phi_X(b) - \Phi_X(a) = \int_a^b \varphi(x)\,dx = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx$$



$$\Phi(-x_0) = 1 - \Phi(x_0) \quad \forall x_0 > 0$$

$$\Phi(-x_0) = 1 - \Phi(x_0) \quad \forall x_0 < 0$$

$$\Phi(0) = 1/2$$

In order to compute X is std. normal.

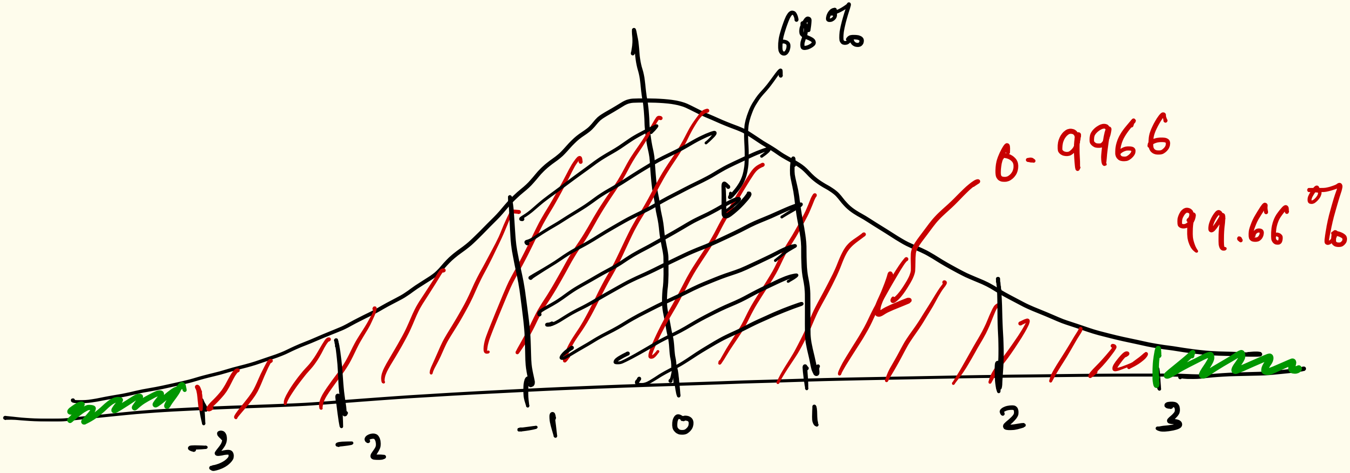
$$P(a \le X \le b) = \Phi(b) - \Phi(a)$$

It is enough to know values of Φ at every real nos. a & b.

Symmetry of X helps!!

It is enough to know the values of Φ only for positive real nos.!!

The values of Φ for positive real numbers are stored in standard normal tables!!



X is std. normal.

$$
\begin{aligned}
P(-3 < X < 3) &= \Phi(3) - \Phi(-3) \\
&= \Phi(3) - (1 - \Phi(3)) \\
&= 2\Phi(3) - 1 \\
&= 2(0.9983) - 1 = 0.9966
\end{aligned}
$$

$P(-1 \le X \le 1) = 0.6826$

$P(-2 \le X \le 2) = 0.95$

$P(X \geq -0.6) = 1 - \Phi(-0.6) = 0.725$

$P(X \le 2.9) = 0.9981$

X is std. normal.