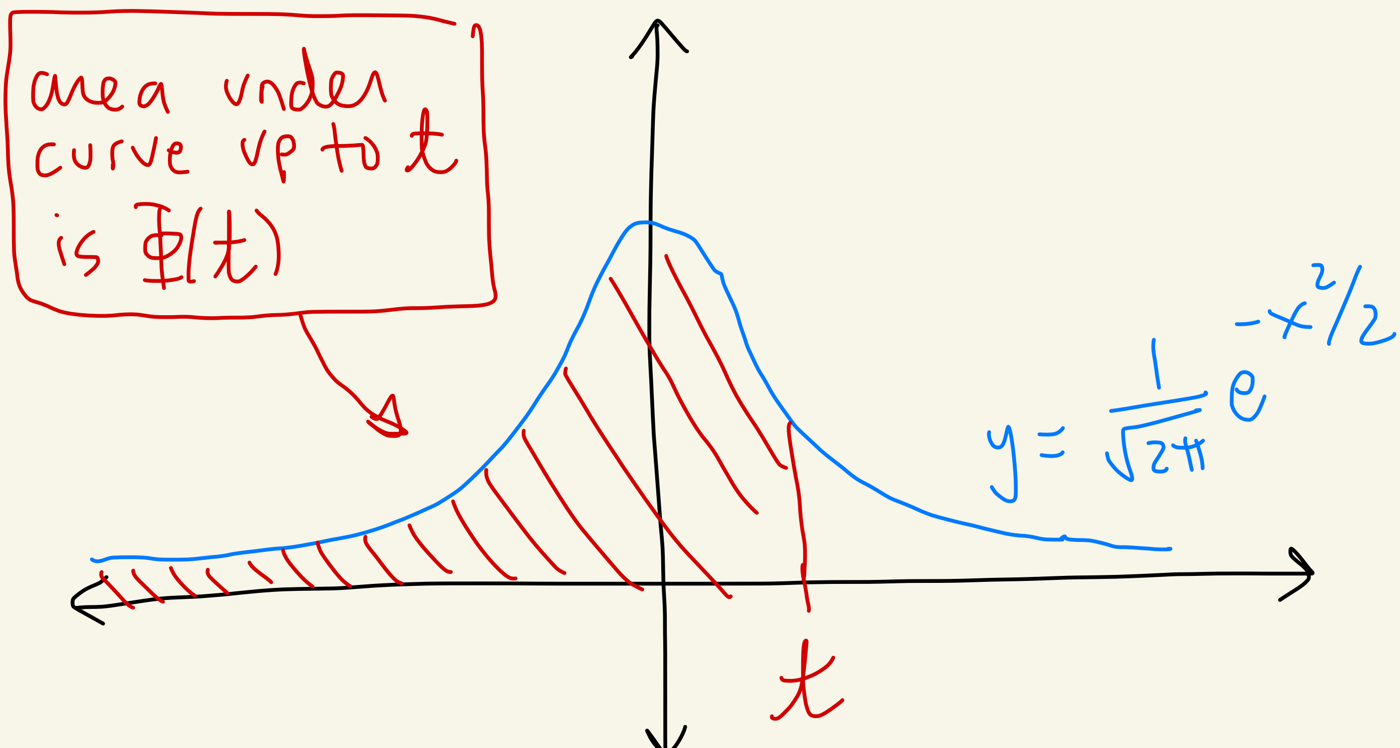
# Topic 7 - Normal approximation to binomial random variables

**Def.** Let

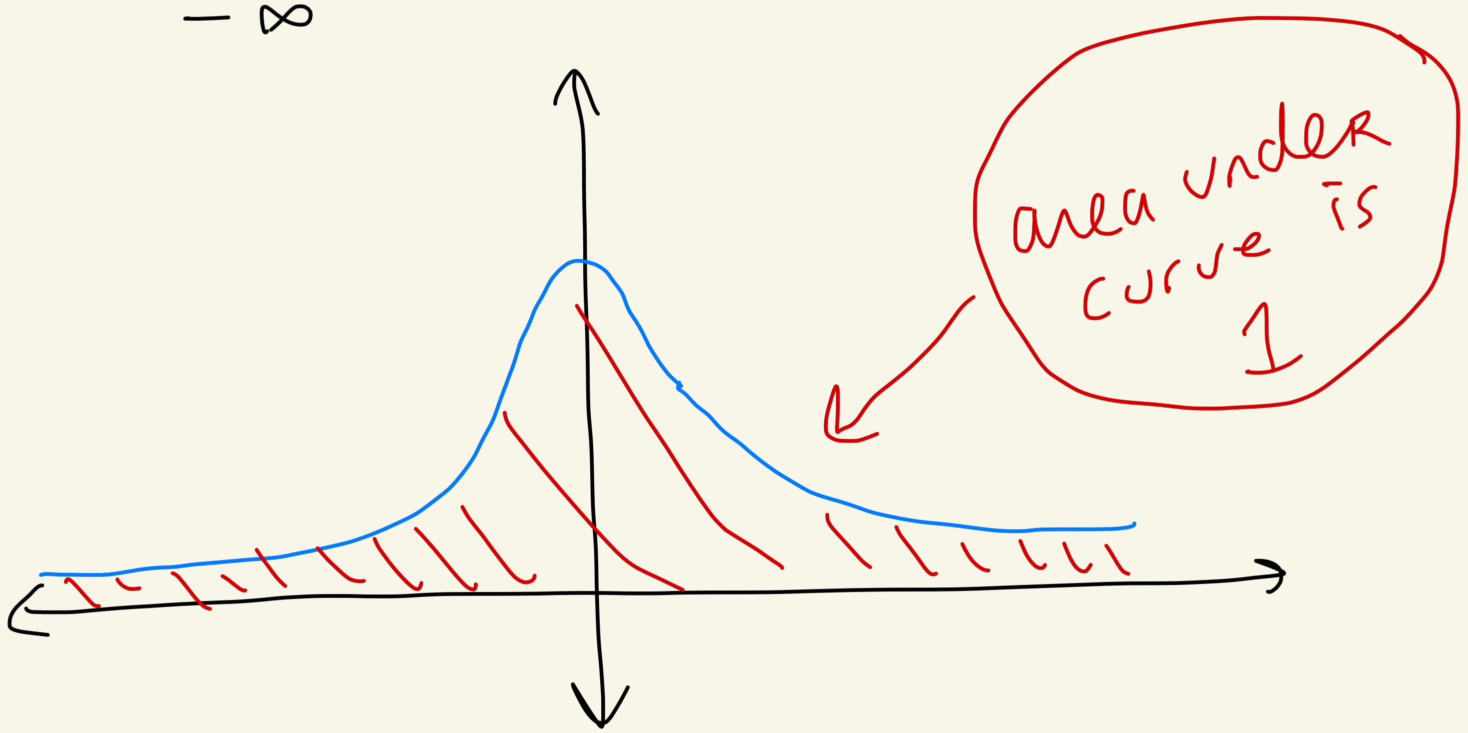
$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2}\, dx$$




$\Phi$ is called the probability density function of the standard normal random variable (topic 8)

In topic 8, we will see that

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} \, dx = 1$$

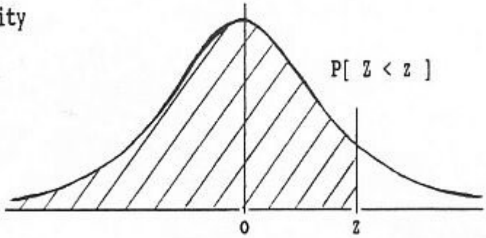


# STANDARD STATISTICAL TABLES

## 1. Areas under the Normal Distribution

The table gives the cumulative probability up to the standardised normal value z
i.e.

$$P[\ Z < z\ ] = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}Z^2)\ dZ$$

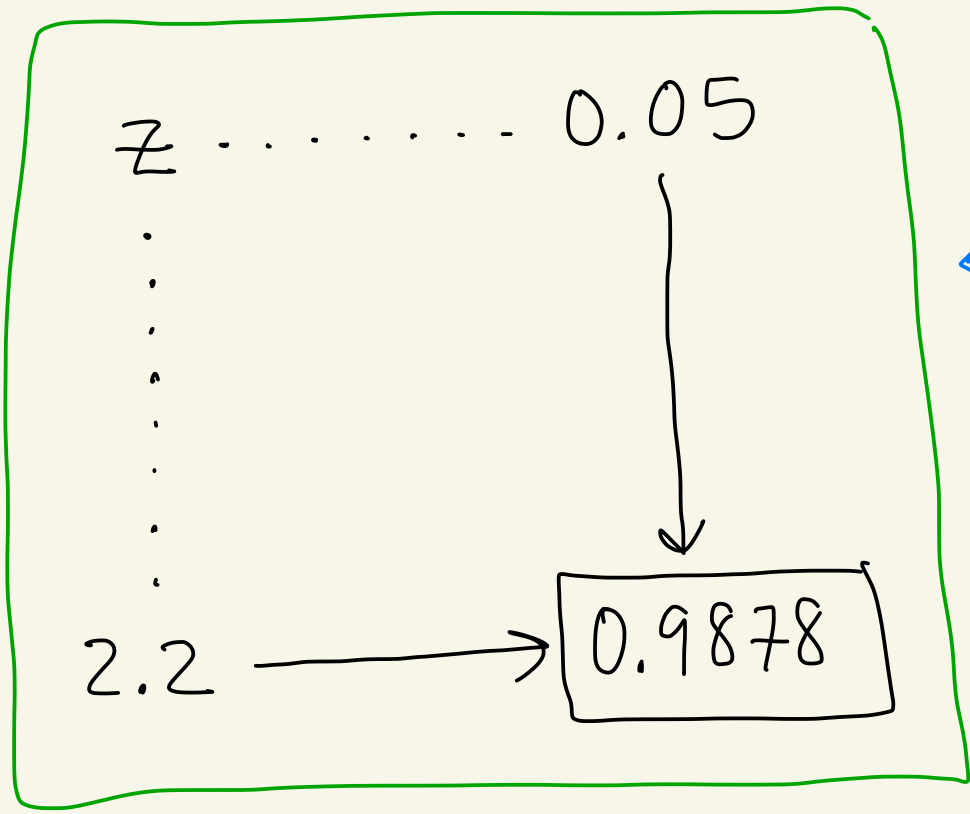



| z | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5159 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| 0.1 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| 0.2 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| 0.3 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| 0.4 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| 0.5 | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |
| 0.6 | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | 0.7422 | 0.7454 | 0.7486 | 0.7517 | 0.7549 |
| 0.7 | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7704 | 0.7734 | 0.7764 | 0.7794 | 0.7823 | 0.7854 |
| 0.8 | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | 0.8023 | 0.8051 | 0.8078 | 0.8106 | 0.8133 |
| 0.9 | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | 0.8289 | 0.8315 | 0.8340 | 0.8365 | 0.8389 |
| 1.0 | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.8554 | 0.8577 | 0.8599 | 0.8621 |
| 1.1 | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8804 | 0.8830 |
| 1.2 | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | 0.8944 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| 1.3 | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| 1.4 | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |
| 1.5 | 0.9332 | 0.9345 | 0.9357 | 0.9370 | 0.9382 | 0.9394 | 0.9406 | 0.9418 | 0.9429 | 0.9441 |
| 1.6 | 0.9452 | 0.9463 | 0.9474 | 0.9484 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| 1.7 | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| 1.8 | 0.9641 | 0.9649 | 0.9656 | 0.9664 | 0.9671 | 0.9678 | 0.9686 | 0.9693 | 0.9699 | 0.9706 |
| 1.9 | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9738 | 0.9744 | 0.9750 | 0.9756 | 0.9761 | 0.9767 |
| 2.0 | 0.9773 | 0.9778 | 0.9783 | 0.9788 | 0.9793 | 0.9798 | 0.9803 | 0.9808 | 0.9812 | 0.9817 |
| 2.1 | 0.9821 | 0.9826 | 0.9830 | 0.9834 | 0.9838 | 0.9842 | 0.9846 | 0.9850 | 0.9854 | 0.9857 |
| 2.2 | 0.9861 | 0.9865 | 0.9868 | 0.9871 | 0.9874 | 0.9878 | 0.9881 | 0.9884 | 0.9887 | 0.9890 |
| 2.3 | 0.9893 | 0.9896 | 0.9898 | 0.9901 | 0.9904 | 0.9906 | 0.9909 | 0.9911 | 0.9913 | 0.9916 |
| 2.4 | 0.9918 | 0.9920 | 0.9922 | 0.9924 | 0.9927 | 0.9929 | 0.9931 | 0.9932 | 0.9934 | 0.9936 |
| 2.5 | 0.9938 | 0.9940 | 0.9941 | 0.9943 | 0.9945 | 0.9946 | 0.9948 | 0.9949 | 0.9951 | 0.9952 |
| 2.6 | 0.9953 | 0.9955 | 0.9956 | 0.9957 | 0.9959 | 0.9960 | 0.9961 | 0.9962 | 0.9963 | 0.9964 |
| 2.7 | 0.9965 | 0.9966 | 0.9967 | 0.9968 | 0.9969 | 0.9970 | 0.9971 | 0.9972 | 0.9973 | 0.9974 |
| 2.8 | 0.9974 | 0.9975 | 0.9976 | 0.9977 | 0.9977 | 0.9978 | 0.9979 | 0.9980 | 0.9980 | 0.9981 |
| 2.9 | 0.9981 | 0.9982 | 0.9982 | 0.9983 | 0.9984 | 0.9984 | 0.9985 | 0.9985 | 0.9986 | 0.9986 |

| z | 3.00 | 3.10 | 3.20 | 3.30 | 3.40 | 3.50 | 3.60 | 3.70 | 3.80 | 3.90 |
|---|---|---|---|---|---|---|---|---|---|---|
| P | 0.9986 | 0.9990 | 0.9993 | 0.9995 | 0.9997 | 0.9998 | 0.9998 | 0.9999 | 0.9999 | 1.0000 |

Let's see how to calculate $\Phi(t)$ when $t \geq 0$

**Ex:** Calculate $\Phi(2.25)$




$2.25 = 2.2 + 0.05$

So, $\Phi(2.25) \approx 0.9878$

**Ex:** Calculate $\Phi(1.36)$

| $z$ | $\cdots$ | $0.06$ |
|---|---|---|
| $\vdots$ | | $\downarrow$ |
| $1.3$ | $\longrightarrow$ | $\boxed{0.9131}$ |

So, $\Phi(1.36) \approx 0.9131$

---

**Ex:** Calculate $\Phi(0.4)$

| $z$ | $0.00$ |
|---|---|
| $\vdots$ | $\downarrow$ |
| $0.4 \rightarrow$ | $\boxed{0.6554}$ |

So,

$\Phi(0.4) \approx 0.6554$

In the table they have

$$\Phi(3.9) \approx 1$$

but its smaller than 1.
Would need a better table to get a better approximation.

If $t \geq 3.9$ you could approx.

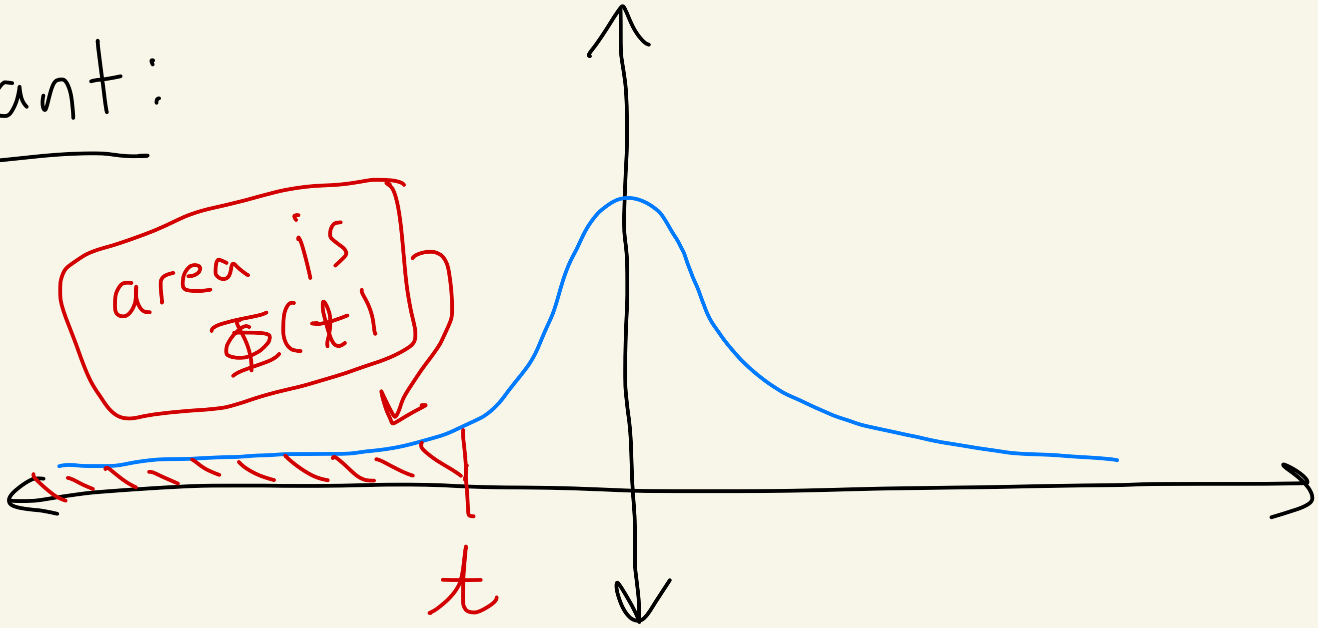
$$\Phi(t) \approx 1$$

It's close to but less than 1.

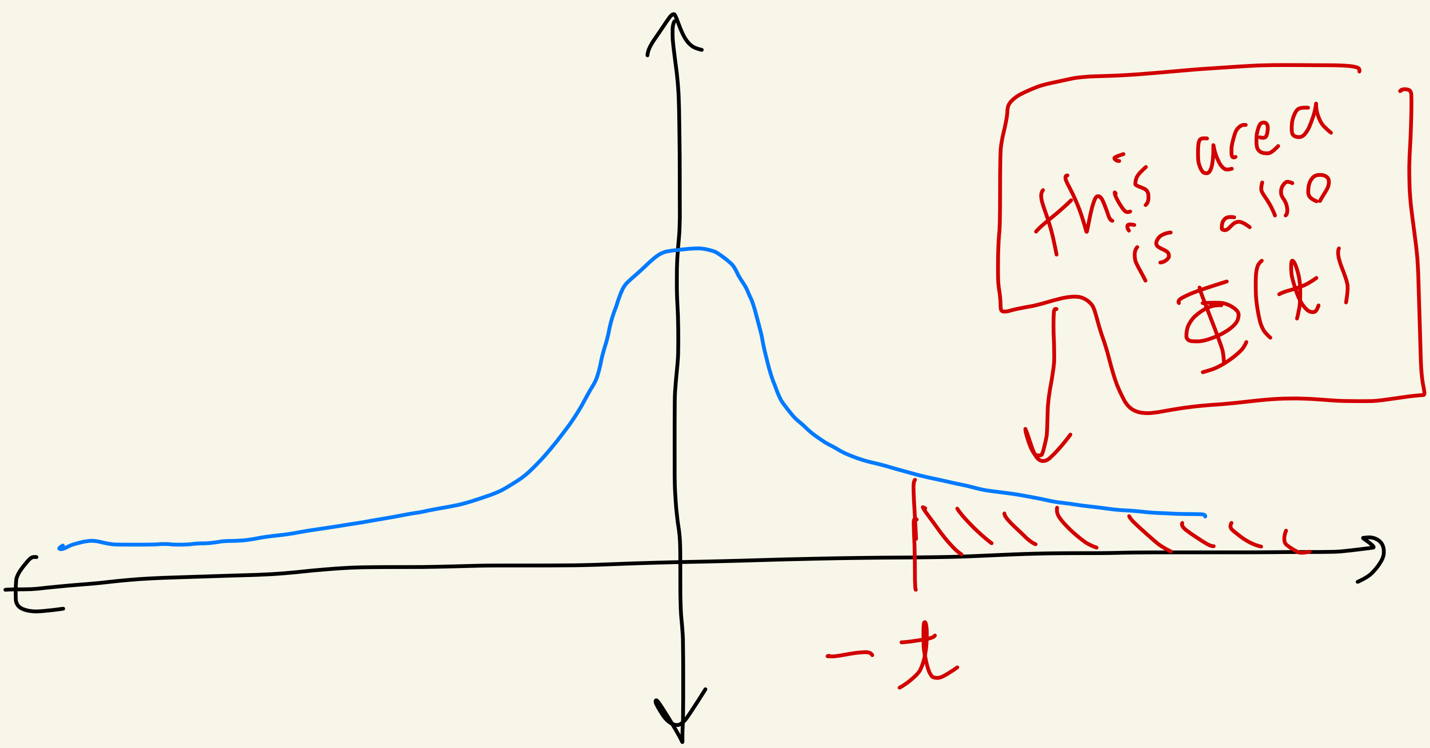
---

# How to calculate $\Phi(t)$ when $t<0$

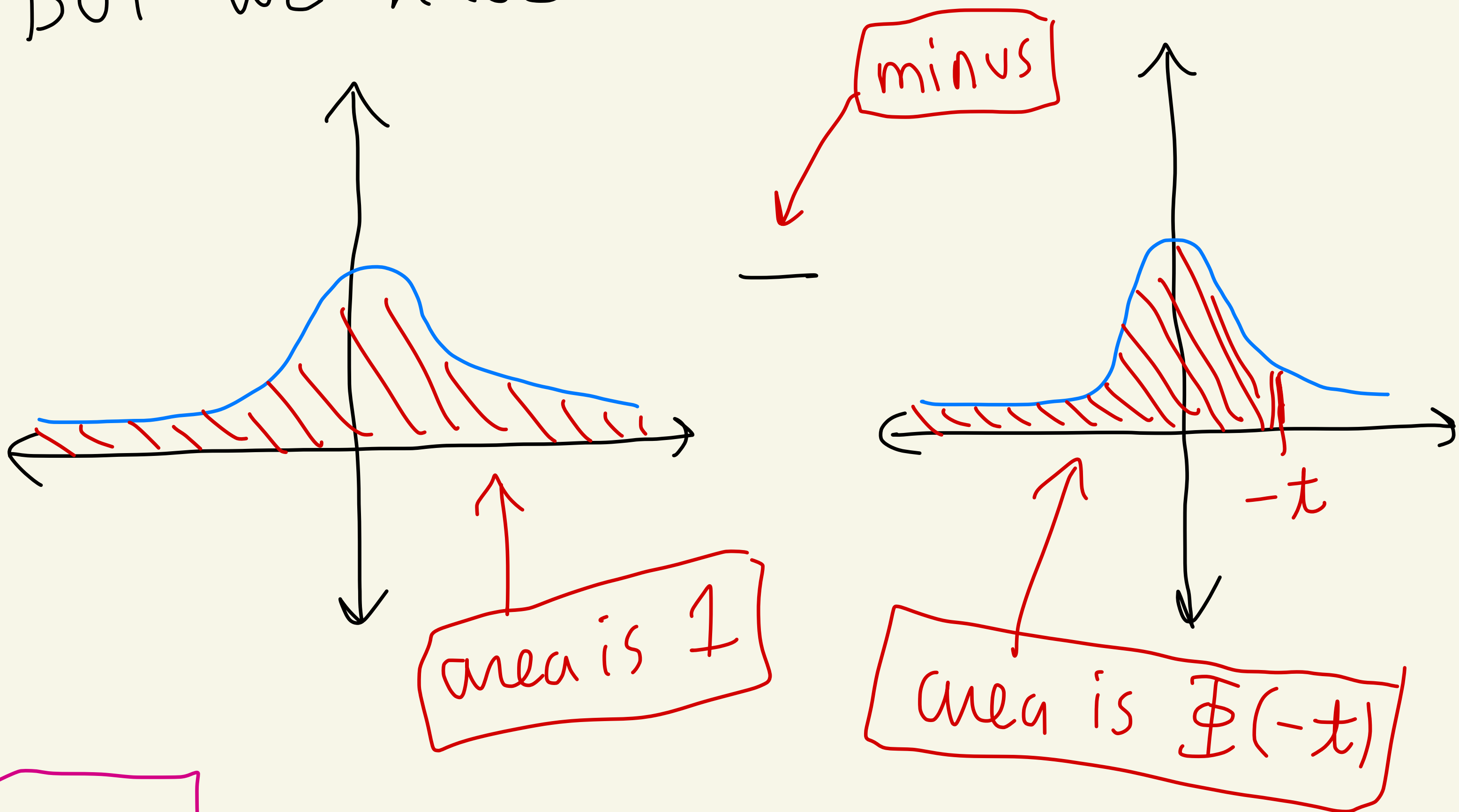
<u>want:</u>




by symmetry $\left(\text{since } \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \text{ is an even function}\right)$



But we have the above area is




So,

$$\boxed{\Phi(t) = 1 - \Phi(-t) \quad \text{if } t<0}$$

<u>Ex:</u>

$$\Phi(-2.68) = 1 - \Phi(2.68)$$

$$\approx 1 - 0.9963 \approx \boxed{0.0037}$$

**Theorem:** (De Moivre – Laplace Theorem)

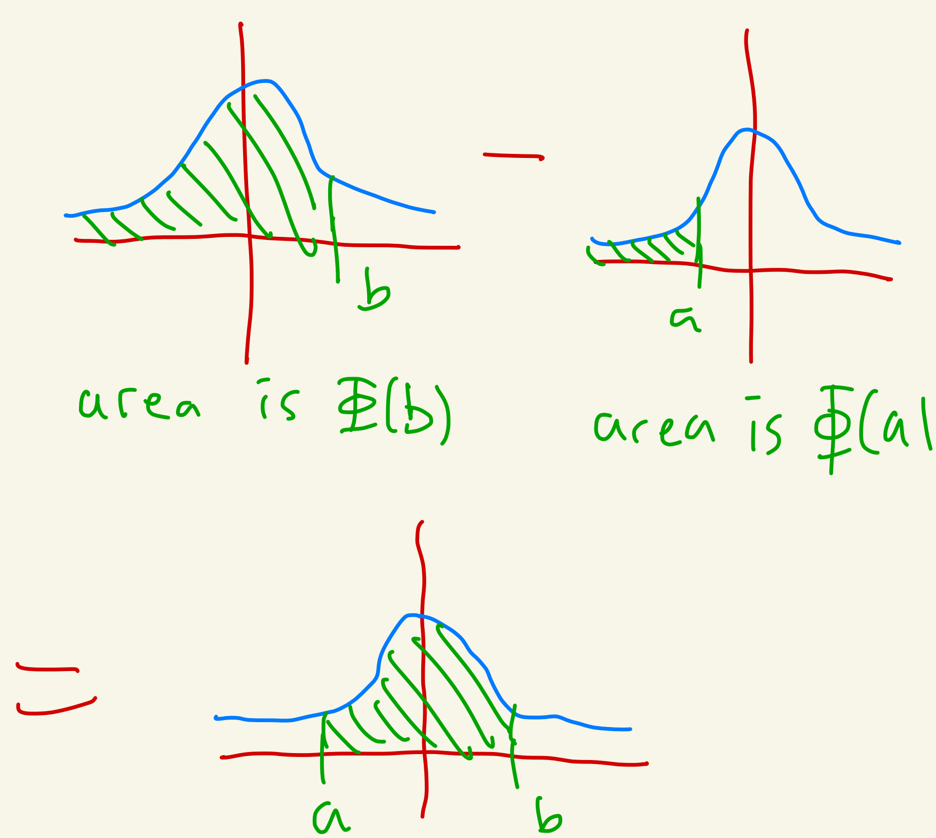
Let $X$ be a binomial random variable with parameters $n$ and $p$. Then for any real numbers $a$ and $b$ with $a<b$ we have that

$$\lim_{n\to\infty} P\left(a \le \frac{X-np}{\sqrt{np(1-p)}} \le b\right) = \Phi(b) - \Phi(a)$$

$$= \frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}\,dx$$

$E[X] = np$

$\sigma_X = \sqrt{np(1-p)}$



You can also do:

$$\lim_{n\to\infty} P\left(\frac{X-np}{\sqrt{np(1-p)}} \le b\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{b} e^{-x^2/2}$$

"$a=-\infty$"

$$= \Phi(b)$$

---

<u>Ex:</u> Suppose we flip a coin 10,000 times. Let $X$ be the number of heads that occur.

Approximate the probability that $5000 \le X \le 5002$.

---

Here $X$ is a binomial random variable with $n = 10{,}000$ and $p = \frac{1}{2}$.

So, $np = 5000$ and $\sqrt{np(1-p)} = \sqrt{2500}$
$= 50$

Thus,

$$P(5000 \le X \le 5002)$$

$$= P\left( \underbrace{\frac{5000-5000}{50}}_{0} \le \overbrace{\frac{X-5000}{50}}^{\frac{X-np}{\sqrt{np(1-p)}}} \le \underbrace{\frac{5002-5000}{50}}_{\approx 0.04} \right)$$

$$\approx P\left(0 \le \frac{X-5000}{50} \le 0.04\right)$$

$$\approx \Phi(0.04) - \Phi(0) \approx 0.5159 - 0.5$$

n = 10,000 is a big #

DeMoivre Laplace

$$\approx 0.0159$$

$$\approx \boxed{1.59\%}$$

Ex: Suppose you flip a coin 40 times. Let $X$ be the number of heads.

Approximate $P(X=20)$.

$$n=40, \quad p=\frac{1}{2} \quad \Big\} \quad np=20, \quad \sqrt{np(1-p)}=\sqrt{10}$$

We have:

$$P(X=20)=P(19.5\le X\le 20.5)$$

$X$ can only be whole # values

$$=P\left(\frac{19.5-20}{\sqrt{10}}\le \underbrace{\frac{X-20}{\sqrt{10}}}_{\frac{X-np}{\sqrt{np(1-p)}}}\le\frac{20.5-20}{\sqrt{10}}\right)$$

$$\approx P\left(-0.16\le\frac{X-20}{\sqrt{10}}\le 0.16\right)$$

DeMoivre Laplace even though $n=40$ is small

$$\approx \Phi(0.16)-\Phi(-0.16)$$

$$= \Phi(0.16) - [1 - \Phi(0.16)]$$

$$\boxed{\Phi(-t) = 1 - \Phi(t)}$$

$$= 2\Phi(0.16) - 1$$

$$\approx 2[0.5636] - 1$$

$$\approx \boxed{0.1272} \approx \boxed{12.72\%}$$

Is this accurate? Yes!

$$P(X = 20) = \binom{40}{20} \cdot \left(\frac{1}{2}\right)^{20} \left(1 - \frac{1}{2}\right)^{40-20}$$

$$= \frac{137,846,528,820}{1,099,511,627,776}$$

$$\approx \boxed{0.125371} \approx \boxed{12.54\%}$$