

IE-231 Normal Distribution Tutorial

Solving normal distribution problems is quite easy if you remember some simple items.

Item 1: Standard Normal Distribution and z-table

Standard normal distribution is special case of normal distribution with mean (μ) of 0 and standard deviation (σ) of 1. The cumulative distribution function (CDF) values ($P(X \leq x)$) of std normal distribution are given in the z-table.

Note: CDF of the standard normal is also denoted with $\Phi(x) = P(X \leq x)$. Note 2: For those of you who did not grasp the concept of CDF; it is basically the probability that the outcome of the event is less than the value specified. For instance $\Phi(0)$

Item 2: Convert and normal process to standard normal.

Suppose you have a process (say, you pour cream on cakes) with mean 50 and standard deviation 5 and you want to know the probability of getting less than 52. Follow this formula.

$$\Phi\left(\frac{X - \mu}{\sigma}\right)$$

So in our case, $\Phi\left(\frac{52 - 50}{5}\right) = \Phi(0.4)$.

Item 3: How to read z-table.

In the z-table sum of row and column names gives us the value (x) and their corresponding value gives us the probability $\Phi(x)$. In our example $\Phi(0.4) = \Phi(\text{row} + \text{column}) = \Phi(0.4 + 0.00) = 0.6554$

Standard Normal Cumulative Probability Table

Cumulative probabilities for POSITIVE z-values are shown in the following table:

| 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 | 0.5160 | 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 |



The notation for getting the (quantile) value is $\Phi^{-1}(0.6554) = 0.4$.

Item 4: Symmetry

The standard normal distribution is symmetric around 0. It gives us the following useful property $\Phi(-x) = 1 - \Phi(x)$. See $\Phi(-0.4) = 1 - 0.6554 = 0.3446$.

| | | | | | | | | | | |
|-------------|---------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| -0.4 | 0.3446 | 0.3409 | 0.3372 | 0.3336 | 0.3300 | 0.3264 | 0.3228 | 0.3192 | 0.3156 | 0.3121 |
| -0.3 | 0.3821 | 0.3783 | 0.3745 | 0.3707 | 0.3669 | 0.3632 | 0.3594 | 0.3557 | 0.3520 | 0.3483 |
| -0.2 | 0.4207 | 0.4168 | 0.4129 | 0.4090 | 0.4052 | 0.4013 | 0.3974 | 0.3936 | 0.3897 | 0.3859 |
| -0.1 | 0.4602 | 0.4562 | 0.4522 | 0.4483 | 0.4443 | 0.4404 | 0.4364 | 0.4325 | 0.4286 | 0.4247 |
| 0.0 | 0.5000 | 0.4960 | 0.4920 | 0.4880 | 0.4840 | 0.4801 | 0.4761 | 0.4721 | 0.4681 | 0.4641 |

Item 5: Understand intervals.

Due to the nature of normal distribution events mainly occur around the mean. That's why central intervals are given with some probability instead of from the bottom to the end. For instance suppose we would like to know the probability around ± 9.8 liters around the mean.

$$P(50 - 9.8 < X < 50 + 9.8) = P(40.2 < X < 59.8) = P(X < 59.8, X > 40.2) = P(X < 59.8) - P(X < 40.2)$$

Convert to standard normal.

$$P(X < (59.8 - 50)/5) - P(X < (40.2 - 50)/5) = \Phi(1.96) - \Phi(-1.96)$$

We know that $\Phi(-x) = 1 - \Phi(x)$, so $\Phi(-1.96) = 1 - \Phi(1.96)$.

$$\Phi(1.96) - \Phi(-1.96) = \Phi(1.96) - (1 - \Phi(1.96)) = 2 * \Phi(1.96) - 1$$

Check the z-table

| z | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|------------|-------------|-------------|-------------|-------------|-------------|-------------|---------------|-------------|-------------|-------------|
| 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 |

$\Phi(1.96)$ is 0.975. So $2 * 0.975 - 1 = 0.95$.

Item 6: Moving the mean

It is possible to desire an increased level of quality. For instance, if we would like to have 90% of the cakes have at least 50 grams of cream instead of 50 grams just being the mean. So $P(X > 50) = 1 - P(X < 50) = 0.9$, then $P(X < 50) = 0.1$. You have the same standard deviation of 5 but you need a new mean.

$$\Phi\left(\frac{50 - \mu^*}{5}\right) = 0.1$$

We should invert the standard normal variate $\Phi^{-1}(0.1)$. Let's check the z-table for the values corresponding to 0.1.

| z | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| -1.4 | 0.0808 | 0.0793 | 0.0778 | 0.0764 | 0.0749 | 0.0735 | 0.0721 | 0.0708 | 0.0694 | 0.0681 |
| -1.3 | 0.0968 | 0.0951 | 0.0934 | 0.0918 | 0.0901 | 0.0885 | 0.0869 | 0.0853 | 0.0838 | 0.0823 |
| -1.2 | 0.1151 | 0.1131 | 0.1112 | 0.1093 | 0.1075 | 0.1056 | 0.1038 | 0.1020 | 0.1003 | 0.0985 |
| -1.1 | 0.1357 | 0.1335 | 0.1314 | 0.1292 | 0.1271 | 0.1251 | 0.1230 | 0.1210 | 0.1190 | 0.1170 |
| -1.0 | 0.1587 | 0.1562 | 0.1539 | 0.1515 | 0.1492 | 0.1469 | 0.1446 | 0.1423 | 0.1401 | 0.1379 |

Well, closest value is 0.1003 so we go with it (usually linear interpolation is done, but we'll skip that). The corresponding value is -1.28.

$$\begin{aligned} \frac{50 - \mu^*}{5} &= -1.28 \\ 50 - \mu^* &= -1.28 * 5 = -6.4 \\ \mu^* &= 50 + 6.4 = 56.4 \end{aligned}$$

So we should align ourself to pour 56.4 grams of cream per cake on average in order to have 90% of the cakes with 50 grams or more cream.

Conclusion

By the end of this tutorial you should know your way around the normal distribution problems. It is not hard but you might need some practice till you get adept on the topic.