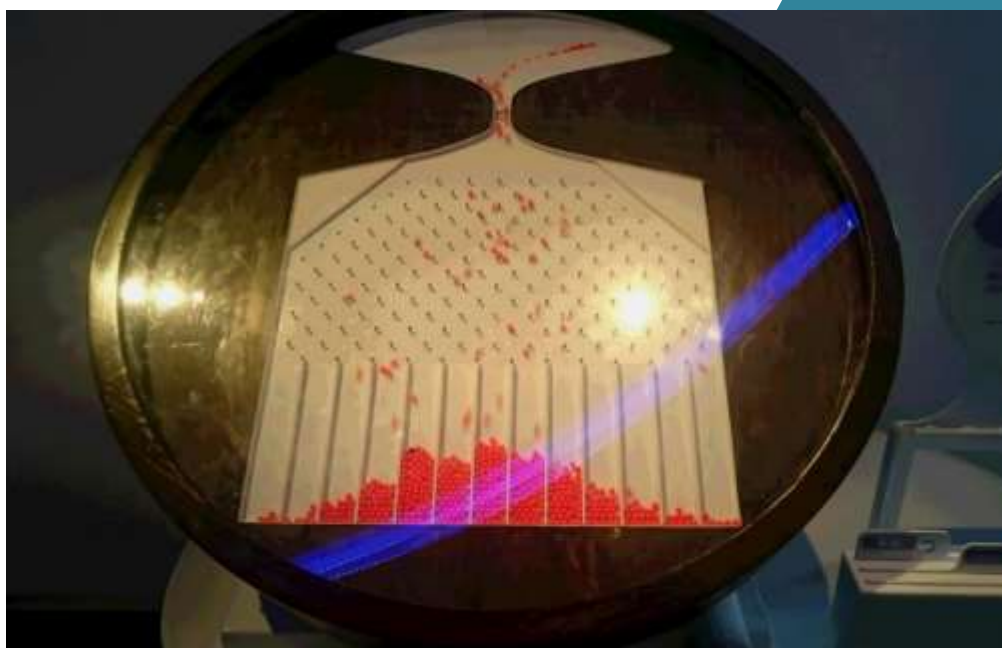


## 2 日目

# 標準正規分布



東京のある博物館には、写真のような装置がありました（現在はこの博物館は閉館になっています。残念ながら）。

これは、上から小さな球を転がすと、途中にある棒に当たりながら落ち、これらの球が下に集まるようになっています。

球の分布は、上のような山型の形になります。このような分布は、わたしたちの身の回りの現象の中にもたくさんあります。

身の回りにあるいろいろなデータの分布のしかたがどのようなになっているか、調べてみましょう。

## 2 日目 標準正規分布

# 1 ヒストグラムを細かくすると…?

ねらい 連続的な値をとる確率変数の確率分布を考えてみよう。

### 1 連続的な確率変数

右の表は、ある学年 100 名の身長を 5 cm 単位で集計した表です。

例えば、身長が 165 cm 以上 170 cm 未満の生徒は、全体の 24% いると見ることができますね。

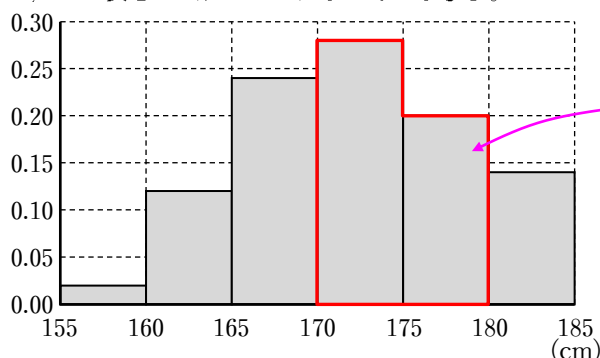
で、この表の見方を変えてみましょう。

**100 人の生徒から無作為に 1 人選んだとき、その生徒の身長  $X$  が 165 cm 以上 170 cm 未満の階級値 167.5 である確率が 24% と見るわけです。**

つまり、 $X$  は階級値の値をとる確率変数と考えることができ、その確率分布は相対度数の分布と一致します。

さて、この表をヒストグラムにしてみましょう。

| 階級(cm)          | 度数(人) | 相対度数 |
|-----------------|-------|------|
| 155 以上 ~ 160 未満 | 2     | 0.02 |
| 160 ~ 165       | 12    | 0.12 |
| 165 ~ 170       | 24    | 0.24 |
| 170 ~ 175       | 28    | 0.28 |
| 175 ~ 180       | 20    | 0.20 |
| 180 ~ 185       | 14    | 0.14 |
| 計               | 100   | 1.00 |



各柱の横の長さを 1 と考えたとき、この赤枠の面積は、身長が 170 cm 以上 180 cm 未満である確率と考えられます。

そうすると、無作為に選んだ 1 人の生徒の身長がある階級に属する確率は、**すべての柱の面積の合計を 1 と考えたときの、その階級の柱の面積で表されます。**

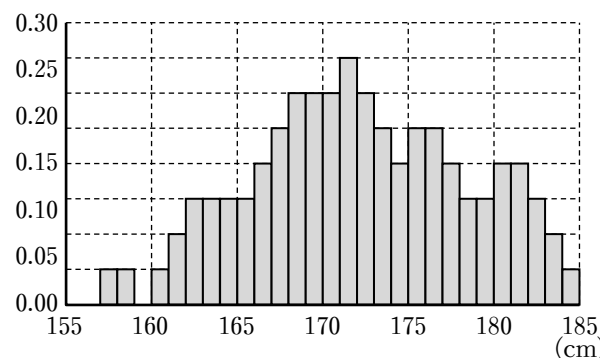
例えば、身長が 170 cm 以上 180 cm 未満となる確率  $P(170 \leq X < 180)$  は、上の赤い枠で示した部分の面積と一致します。

今は階級の幅を 5 cm にしてヒストグラムをかきましたが、これを 1 cm にしたら何が起きるでしょうか？

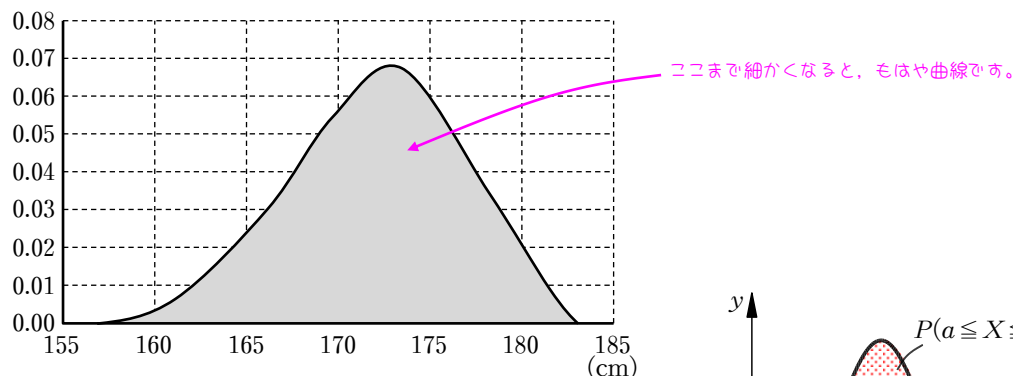
……当然のことですが、右の図のように、柱がだいぶ細くなります。

と同時に、「山の形」が見えてきました。

そして……。

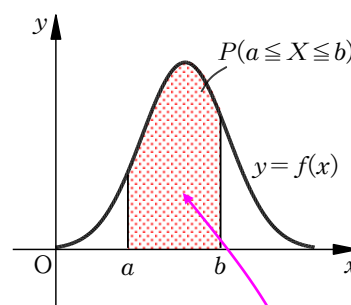


で、さらに細かく細か〜く刻んでいくと、やがてこんな感じになるはずです。



そこで、一般に、連続的な値をとる確率変数  $X$  の確率分布を考える場合には、右の図のように  $X$  に 1 つの曲線  $y=f(x)$  を対応させ、 $a \leq X \leq b$  となる確率が、図の赤い部分の面積で表されるようにします。

このような曲線を、人呼んで**分布曲線**と申します。



### 分布曲線の性質

確率変数  $X$  の分布曲線の方程式を  $y=f(x)$  とすると、関数  $f(x)$  は次の性質を持ちます。

- [1] 常に  $f(x) \geq 0$
- [2]  $X$  が  $a \leq X \leq b$  となる確率は、次の定積分で求めることができる。

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$

- [3]  $X$  のとり得る値の範囲が  $m \leq X \leq M$  のとき

$$P(m \leq X \leq M) = \int_m^M f(x) dx = 1$$

確率の総和は 1 であってことを言っているだけです。

ついでに用語を整理しておくと、これまで扱ってきたとびとびの値をとる確率変数を**離散型確率変数**、今回のように連続的な値をとる確率変数を**連続型確率変数**といいます。

さらに、関数  $f(x)$  を、**確率密度関数**と呼びます。

### COLUMN : 確率“密度”？

連続型確率変数を考える場合、「 $X$  がある値  $a$  ちょうどになる確率」を考えることに意味はありません。さっきの身長例で言えば、「ある人が（寸分の狂いもなく）160.0000 cm である」なんてことはきっとありませんよね。そうすると、身長が 160.0000 cm である確率は 0 になってしまうんです。

だから、関数  $f(x)$  は  $X=x$  での「確率」そのものではありません。 $X$  がある範囲を示したときに確率が求められるように、つまり、その範囲の幅をかけることで確率が得られるように、確率「密度」という概念を導入したわけです。

もうちょっと突っ込んだ言い方をすると、連続型確率変数  $X$  がある値  $x$  とそれに限りなく近い  $x+dx$  の間の値をとるとき、その確率を

$$f(x) dx$$

と考えるよってことです。

ヒストグラムで、横の幅  $dx$  が限りなく 0 に近い柱の面積を考えています。

## 2

## 平均・分散はどう定義する？

離散型確率変数での平均・分散を考えたように、連続型確率変数でも平均・分散を考えたいと思います。できるだけ自然な拡張となるように定義するにはどうすればよいでしょうか？

離散型確率変数  $X$  の平均は

「( $X$  のとる値)  $\times$  (その値をとる確率)」の和 ←  $X$  が  $x_1, \dots, x_n$  の値をとる確率が  $p_1, \dots, p_n$  であれば

$$E(X) = \sum_{k=1}^n x_k p_k$$

でした。

そこで、連続型確率変数  $X$  の平均も

「( $X$  のとる値)  $\times$  (その値をとる確率)」の和

で定義したいなあというのが人情です。

連続型の場合、 $X = x$  である確率（正確には  $x$  と  $x + dx$  の間の値をとる確率）は、確率密度関数  $f(x)$  を使って

$$f(x) dx$$

で表されたので、 $X$  のとり得る値の範囲が  $m \leq X \leq M$  であれば、平均  $E(X)$  を

$$E(X) = \int_m^M x f(x) dx$$

で定義するのがよさそうです。

じゃあ分散は？

考え方は同じです。離散型の場合と同様に定義します。

つまり、 $E(X) = \mu$  とすると

$$V(X) = \int_m^M (x - \mu)^2 f(x) dx$$

と定義できます。

思い出して！ 平均を  $m$  とすると

$$V(X) = \sum_{k=1}^n (x_k - m)^2 p_k$$

でした。

### ● COLUMN : 分散を求める公式の連続版は？ ●

分散は、 $V(X) = E(X^2) - \{E(X)\}^2$  なる時短テクがありました。連続型確率変数の場合も同じ公式があって

$$V(X) = \int_m^M x^2 f(x) dx - \left\{ \int_m^M x f(x) dx \right\}^2$$

と計算することもできます。

まとめておきましょう！！

### 連続型確率変数の平均・分散

確率変数  $X$  のとり得る値の範囲が  $m \leq X \leq M$  で、その確率密度関数が  $f(x)$  のとき、 $X$  の平均  $E(X)$  と分散  $V(X)$  は、 $E(X) = \mu$  として

$$E(X) = \int_m^M x f(x) dx$$

$$V(X) = \int_m^M (x - \mu)^2 f(x) dx$$

で定義されます。このとき

$$V(X) = \int_m^M x^2 f(x) dx - \{E(X)\}^2$$

が成り立ちます。

## 2 日目 標準正規分布

# 2 最重要！正規分布とは!?

ねらい 連続型確率変数の確率分布のうち、代表的な正規分布について学ぼう。

### 1 正規分布って何？

本題に入る前に、まずは復習からいきましょう。

コインを 100 回投げたとき表が出る回数を  $X$  とするとき、 $X$  の平均  $E(X)$  はいくつになるか考えてみます。これは「表が出るか」、「表が出ない（裏が出る）か」の 2 通りの結果しか起こらないので、表が出る回数  $X$  は二項分布  $B(100, \frac{1}{2})$  に従うことになります。

つまり、 $X$  の平均  $E(X)$  は

$$E(X) = 100 \times \frac{1}{2} = 50$$

確率変数  $X$  が  $B(n, p)$  に従うとき、

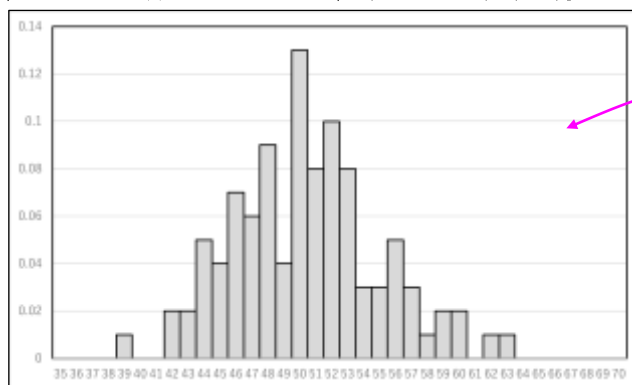
$$E(X) = np$$

でした。

と求められますね。

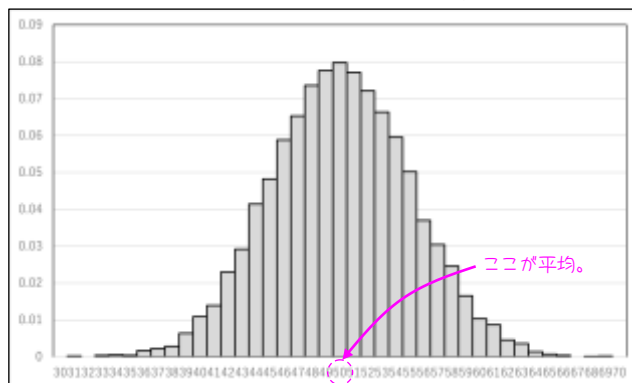
では、「コインを 100 回投げる」という実験を、何セットか繰り返して横軸に表が出た回数、縦軸に相対度数をとったヒストグラムを作ってみます。

試みに、100 セット繰り返してみると、下のような結果が得られました。



実際にやるのは大変なので、Excel 先輩に頑張ってもらいました。

で、その 100 倍の 10000 セット繰り返してみると、下のような結果が得られました。



100 セットのときよりも、きれいな山型になりました。このように、セットを増やしていくと、やがてこの山型は、平均付近が大きく、平均から離れれば離れるほど確率が減っていくような形になります。

この山のような分布こそがまさに、**正規分布**なのです！！

## 2

## 標準正規分布

正規分布のグラフの形（分布の仕方）は、確率変数  $X$  の平均  $m$  と標準偏差  $\sigma$  の 2 つによって決まります。平均  $m$ 、標準偏差  $\sigma$ （分散  $\sigma^2$ ）の正規分布を、記号  $N(m, \sigma^2)$  で表します。

で、その特別なやつが**標準正規分布**です。

## 標準正規分布

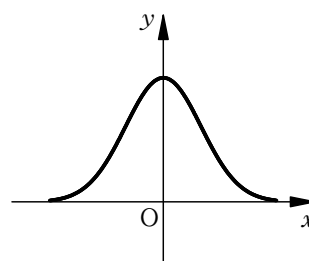
正規分布  $N(m, \sigma^2)$  のうち、平均が 0 で標準偏差が 1 となるものを**標準正規分布**といって、記号  $N(0, 1)$  で表します。

標準正規分布は確率密度関数の 1 つで

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

で表され、右のような分布曲線になります。

この式は覚えなくて OK！ 一般的な話を一応あると……  
正規分布  $N(m, \sigma^2)$  の確率密度関数は  $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$  だよ。



標準正規分布

で、p.21 で説明したように、確率密度関数  $f(x)$  で表される確率変数  $X$  が  $a \leq X \leq b$  になる確率は

$$\int_a^b f(x) dx$$

を計算すれば求められるわけですが……

標準正規分布  $N(0, 1)$  に従う確率変数  $X$  が  $a \leq X \leq b$  となる確率

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

は高校数学を逸脱した積分計算が必要になってしまいます。

そこで！ 大学入試の問題では、上の積分の計算結果があらかじめ表で与えられているんです。

それが次のページに示した**正規分布表**です！！

正規分布表は、標準正規分布に従う確率変数  $X$  が  $0 \leq X \leq z_0$  になる確率が小数で与えられています。

ここでは、正規分布表を読み取る上でもっとも基本的な

$$P(0 \leq X \leq z_0)$$

の求め方について

$$P(0 \leq X \leq 0.56)$$

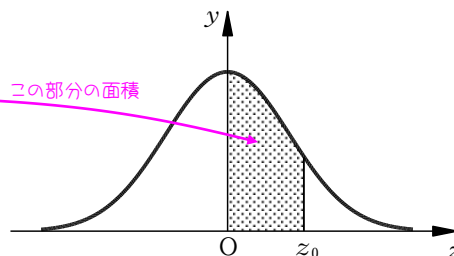
を例に説明しましょう。

この表の一番左の列（縦の並び）は、 $z_0$  の整数部分と小数第 1 位を表します。また、一番上の行（横の並び）は  $z_0$  の小数第 2 位を表します。

つまり、 $z_0 = 0.56$  の場合、縦の並びは 0.5 の部分を、横の並びは 0.06 の部分を見ればよいということになり、

$$P(0 \leq X \leq 0.56) = 0.2123$$

であることがわかります。



小数第 1 位以上  
 $z_0 = 0.56$

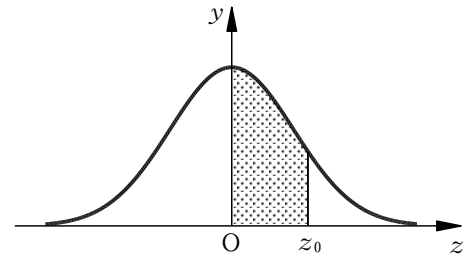
| $z_0$ | 0.00   | 0.01   | 0.02   | 0.03   | 0.04   | 0.05   | 0.06   | 0.07   | 0.08   | 0.09   |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.0   | 0.0000 | 0.0040 | 0.0080 | 0.0120 | 0.0160 | 0.0199 | 0.0239 | 0.0279 | 0.0319 | 0.0359 |
| 0.1   | 0.0398 | 0.0438 | 0.0478 | 0.0517 | 0.0557 | 0.0596 | 0.0636 | 0.0675 | 0.0714 | 0.0753 |
| 0.2   | 0.0793 | 0.0832 | 0.0871 | 0.0910 | 0.0948 | 0.0987 | 0.1026 | 0.1064 | 0.1103 | 0.1141 |
| 0.3   | 0.1179 | 0.1217 | 0.1255 | 0.1293 | 0.1331 | 0.1368 | 0.1406 | 0.1443 | 0.1480 | 0.1517 |
| 0.4   | 0.1554 | 0.1591 | 0.1628 | 0.1664 | 0.1700 | 0.1736 | 0.1772 | 0.1808 | 0.1844 | 0.1879 |
| 0.5   | 0.1915 | 0.1950 | 0.1985 | 0.2019 | 0.2054 | 0.2088 | 0.2123 | 0.2157 | 0.2190 | 0.2224 |

小数第 2 位  
 $z_0 = 0.56$

p.27 からは、この表のいろいろな使い方を紹介しましょう。

# 正 規 分 布 表

次の表は、標準正規分布の正規分布曲線における右図の灰色部分の面積の値をまとめたものである。



| $z_0$ | 0.00   | 0.01   | 0.02   | 0.03   | 0.04   | 0.05   | 0.06   | 0.07   | 0.08   | 0.09   |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.0   | 0.0000 | 0.0040 | 0.0080 | 0.0120 | 0.0160 | 0.0199 | 0.0239 | 0.0279 | 0.0319 | 0.0359 |
| 0.1   | 0.0398 | 0.0438 | 0.0478 | 0.0517 | 0.0557 | 0.0596 | 0.0636 | 0.0675 | 0.0714 | 0.0753 |
| 0.2   | 0.0793 | 0.0832 | 0.0871 | 0.0910 | 0.0948 | 0.0987 | 0.1026 | 0.1064 | 0.1103 | 0.1141 |
| 0.3   | 0.1179 | 0.1217 | 0.1255 | 0.1293 | 0.1331 | 0.1368 | 0.1406 | 0.1443 | 0.1480 | 0.1517 |
| 0.4   | 0.1554 | 0.1591 | 0.1628 | 0.1664 | 0.1700 | 0.1736 | 0.1772 | 0.1808 | 0.1844 | 0.1879 |
| 0.5   | 0.1915 | 0.1950 | 0.1985 | 0.2019 | 0.2054 | 0.2088 | 0.2123 | 0.2157 | 0.2190 | 0.2224 |
| 0.6   | 0.2257 | 0.2291 | 0.2324 | 0.2357 | 0.2389 | 0.2422 | 0.2454 | 0.2486 | 0.2517 | 0.2549 |
| 0.7   | 0.2580 | 0.2611 | 0.2642 | 0.2673 | 0.2704 | 0.2734 | 0.2764 | 0.2794 | 0.2823 | 0.2852 |
| 0.8   | 0.2881 | 0.2910 | 0.2939 | 0.2967 | 0.2995 | 0.3023 | 0.3051 | 0.3078 | 0.3106 | 0.3133 |
| 0.9   | 0.3159 | 0.3186 | 0.3212 | 0.3238 | 0.3264 | 0.3289 | 0.3315 | 0.3340 | 0.3365 | 0.3389 |
| 1.0   | 0.3413 | 0.3438 | 0.3461 | 0.3485 | 0.3508 | 0.3531 | 0.3554 | 0.3577 | 0.3599 | 0.3621 |
| 1.1   | 0.3643 | 0.3665 | 0.3686 | 0.3708 | 0.3729 | 0.3749 | 0.3770 | 0.3790 | 0.3810 | 0.3830 |
| 1.2   | 0.3849 | 0.3869 | 0.3888 | 0.3907 | 0.3925 | 0.3944 | 0.3962 | 0.3980 | 0.3997 | 0.4015 |
| 1.3   | 0.4032 | 0.4049 | 0.4066 | 0.4082 | 0.4099 | 0.4115 | 0.4131 | 0.4147 | 0.4162 | 0.4177 |
| 1.4   | 0.4192 | 0.4207 | 0.4222 | 0.4236 | 0.4251 | 0.4265 | 0.4279 | 0.4292 | 0.4306 | 0.4319 |
| 1.5   | 0.4332 | 0.4345 | 0.4357 | 0.4370 | 0.4382 | 0.4394 | 0.4406 | 0.4418 | 0.4429 | 0.4441 |
| 1.6   | 0.4452 | 0.4463 | 0.4474 | 0.4484 | 0.4495 | 0.4505 | 0.4515 | 0.4525 | 0.4535 | 0.4545 |
| 1.7   | 0.4554 | 0.4564 | 0.4573 | 0.4582 | 0.4591 | 0.4599 | 0.4608 | 0.4616 | 0.4625 | 0.4633 |
| 1.8   | 0.4641 | 0.4649 | 0.4656 | 0.4664 | 0.4671 | 0.4678 | 0.4686 | 0.4693 | 0.4699 | 0.4706 |
| 1.9   | 0.4713 | 0.4719 | 0.4726 | 0.4732 | 0.4738 | 0.4744 | 0.4750 | 0.4756 | 0.4761 | 0.4767 |
| 2.0   | 0.4772 | 0.4778 | 0.4783 | 0.4788 | 0.4793 | 0.4798 | 0.4803 | 0.4808 | 0.4812 | 0.4817 |
| 2.1   | 0.4821 | 0.4826 | 0.4830 | 0.4834 | 0.4838 | 0.4842 | 0.4846 | 0.4850 | 0.4854 | 0.4857 |
| 2.2   | 0.4861 | 0.4864 | 0.4868 | 0.4871 | 0.4875 | 0.4878 | 0.4881 | 0.4884 | 0.4887 | 0.4890 |
| 2.3   | 0.4893 | 0.4896 | 0.4898 | 0.4901 | 0.4904 | 0.4906 | 0.4909 | 0.4911 | 0.4913 | 0.4916 |
| 2.4   | 0.4918 | 0.4920 | 0.4922 | 0.4925 | 0.4927 | 0.4929 | 0.4931 | 0.4932 | 0.4934 | 0.4936 |
| 2.5   | 0.4938 | 0.4940 | 0.4941 | 0.4943 | 0.4945 | 0.4946 | 0.4948 | 0.4949 | 0.4951 | 0.4952 |
| 2.6   | 0.4953 | 0.4955 | 0.4956 | 0.4957 | 0.4959 | 0.4960 | 0.4961 | 0.4962 | 0.4963 | 0.4964 |
| 2.7   | 0.4965 | 0.4966 | 0.4967 | 0.4968 | 0.4969 | 0.4970 | 0.4971 | 0.4972 | 0.4973 | 0.4974 |
| 2.8   | 0.4974 | 0.4975 | 0.4976 | 0.4977 | 0.4977 | 0.4978 | 0.4979 | 0.4979 | 0.4980 | 0.4981 |
| 2.9   | 0.4981 | 0.4982 | 0.4982 | 0.4983 | 0.4984 | 0.4984 | 0.4985 | 0.4985 | 0.4986 | 0.4986 |
| 3.0   | 0.4987 | 0.4987 | 0.4987 | 0.4988 | 0.4988 | 0.4989 | 0.4989 | 0.4989 | 0.4990 | 0.4990 |

## 3

## 正規分布表の使い方

いくつかのパターンに分けて考えます。

### $-z_0 \leq X \leq 0$ のタイプ

標準正規分布は、 $y$  軸について対称になっています。

ってことは

$$P(-z_0 \leq X \leq 0) = P(0 \leq X \leq z_0)$$

が成り立ちます。

正規分布表に  $z_0$  が負になる場合が書かれていないのは、正の値を見れば  $y$  軸対称を利用して負の場合も求められるからです。

例えば、 $P(-1.25 \leq X \leq 0)$  を求める場合は、下の表から

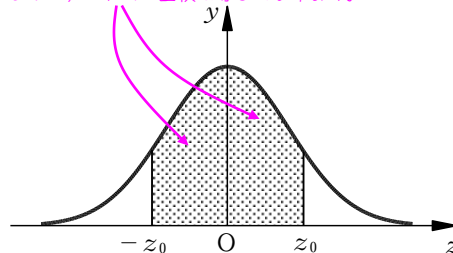
$$P(-1.25 \leq X \leq 0) = P(0 \leq X \leq 1.25)$$

$$= 0.3944$$

と求められます。

| $z_0$ | 0.00   | 0.01   | 0.02   | 0.03   | 0.04   | 0.05   | 0.06   | 0.07   | 0.08   | 0.09   |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.0   | 0.0000 | 0.0040 | 0.0080 | 0.0120 | 0.0160 | 0.0199 | 0.0239 | 0.0279 | 0.0319 | 0.0359 |
| 0.1   | 0.0398 | 0.0438 | 0.0478 | 0.0517 | 0.0557 | 0.0596 | 0.0636 | 0.0675 | 0.0714 | 0.0753 |
| 0.2   | 0.0793 | 0.0832 | 0.0871 | 0.0910 | 0.0948 | 0.0987 | 0.1026 | 0.1064 | 0.1103 | 0.1141 |
| 0.3   | 0.1179 | 0.1217 | 0.1255 | 0.1293 | 0.1331 | 0.1368 | 0.1406 | 0.1443 | 0.1480 | 0.1517 |
| 0.4   | 0.1554 | 0.1591 | 0.1628 | 0.1664 | 0.1700 | 0.1736 | 0.1772 | 0.1808 | 0.1844 | 0.1879 |
| 0.5   | 0.1915 | 0.1950 | 0.1985 | 0.2019 | 0.2054 | 0.2088 | 0.2123 | 0.2157 | 0.2190 | 0.2224 |
| 0.6   | 0.2257 | 0.2291 | 0.2324 | 0.2357 | 0.2389 | 0.2422 | 0.2454 | 0.2486 | 0.2517 | 0.2549 |
| 0.7   | 0.2580 | 0.2611 | 0.2642 | 0.2673 | 0.2704 | 0.2734 | 0.2764 | 0.2794 | 0.2823 | 0.2852 |
| 0.8   | 0.2881 | 0.2910 | 0.2939 | 0.2967 | 0.2995 | 0.3023 | 0.3051 | 0.3078 | 0.3106 | 0.3133 |
| 0.9   | 0.3159 | 0.3186 | 0.3212 | 0.3238 | 0.3264 | 0.3289 | 0.3315 | 0.3340 | 0.3365 | 0.3389 |
| 1.0   | 0.3413 | 0.3438 | 0.3461 | 0.3485 | 0.3508 | 0.3531 | 0.3554 | 0.3577 | 0.3599 | 0.3621 |
| 1.1   | 0.3643 | 0.3665 | 0.3686 | 0.3708 | 0.3729 | 0.3749 | 0.3770 | 0.3790 | 0.3810 | 0.3830 |
| 1.2   | 0.3849 | 0.3869 | 0.3888 | 0.3907 | 0.3925 | 0.3944 | 0.3962 | 0.3980 | 0.3997 | 0.4015 |
| 1.3   | 0.4032 | 0.4049 | 0.4066 | 0.4082 | 0.4099 | 0.4115 | 0.4131 | 0.4147 | 0.4162 | 0.4177 |

標準正規分布は  $y$  軸について対称です。  
よって、これらの面積は等しくなります。



標準正規分布

### $-z_0 \leq X \leq z_1$ のタイプ

確率変数のとる値が負の数から正の数までにわたっている場合です。

このときは、マイナス～0 と 0～プラスの 2 つに分けてそれぞれを求めて、足してあげれば OK です。

つまり

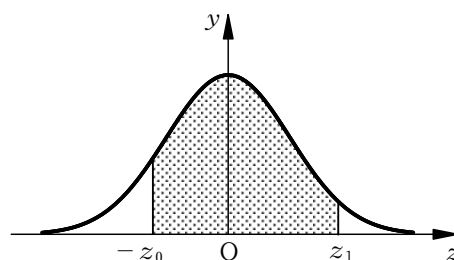
$$\begin{aligned} P(-z_0 \leq X \leq z_1) &= P(-z_0 \leq X \leq 0) + P(0 \leq X \leq z_1) \\ &= P(0 \leq X \leq z_0) + P(0 \leq X \leq z_1) \end{aligned}$$

とすれば求められます！！

例えば

$$\begin{aligned} P(-1.25 \leq X \leq 0.56) &= P(-1.25 \leq X \leq 0) + P(0 \leq X \leq 0.56) \\ &= P(0 \leq X \leq 1.25) + P(0 \leq X \leq 0.56) \\ &= 0.3944 + 0.2123 = 0.6067 \end{aligned}$$

ってな具合です。





### $z_0 \leq X \leq z_1$ のタイプ

ここまで来たらわかると思いますが、今度はひき算パターンです。

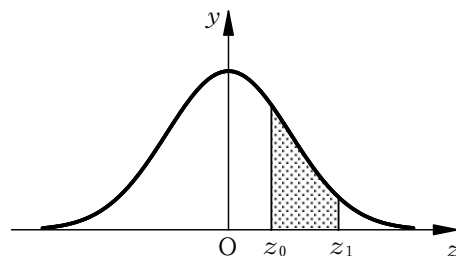
この場合、0 から  $z_1$  までの面積から、0 から  $z_0$  までの面積を引けば求めることができますね。

つまり

$$P(z_0 \leq X \leq z_1) = P(0 \leq X \leq z_1) - P(0 \leq X \leq z_0)$$

とすれば OK です。

$$\begin{aligned} \text{例えば} \quad P(0.56 \leq X \leq 1.25) &= P(0 \leq X \leq 1.25) - P(0 \leq X \leq 0.56) \\ &= 0.3944 - 0.2123 = 0.1821 \quad \text{って感じです。} \end{aligned}$$



### $z_0 \leq X$ のタイプ

最後に、 $X$  がある値以上となる場合です。

標準正規分布の確率密度関数は  $y$  軸について対称なので、0 を境に右半分と左半分の面積は等しくなります。

さらに、全体の面積は 1 なので、左右それぞれの面積は 0.5 になるはずですが。

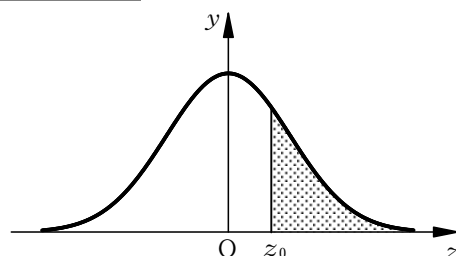
確率の総和は 1 ってヤリよ。

ってなワケで

$$P(X \geq z_0) = 0.5 - P(0 \leq X \leq z_0)$$

とすりゃあ OK です。

$$\begin{aligned} \text{例えば} \quad P(X \geq 1.25) &= 0.5 - P(0 \leq X \leq 1.25) \\ &= 0.5 - 0.3944 = 0.1056 \quad \text{という感じです。} \end{aligned}$$



## 4

### 正規分布から標準正規分布へ

大学入試で登場する正規分布は、平均 0、標準偏差 1 の標準正規分布  $N(0, 1)$  でない場合がほとんどです。が、試験では標準正規分布以外の正規分布表は与えられません。

ってことで、自力で平均  $m$ 、標準偏差  $\sigma$  の正規分布  $N(m, \sigma^2)$  を標準正規分布の形に直さなければなりません。ここでは、その方法を伝授しましょう。

#### 手順1 平均を 0 に補正する

平均  $m$  の分布を平均 0 にするためには、 $x$  から  $m$  を引いてあげる必要があります。

つまり、 $t = x - m$  とすることで、 $t$  の平均を 0 にすることができます。

平均は 0 に補正されているので、割っても平均は 0 のまま変わりません。

#### 手順2 標準偏差を 1 に補正する

標準偏差が  $\sigma$  の確率分布を標準偏差 1 にするために、 $t$  を  $\sigma$  で割ってあげましょう。

つまり、 $z = \frac{t}{\sigma} = \frac{x - m}{\sigma}$  とすることで、 $z$  の標準偏差を 1 にすることができますね。

このように、正規分布に従う確率変数を標準正規分布に従うように変換することを標準化といいます。

#### 標準正規分布

確率変数  $X$  が正規分布  $N(m, \sigma^2)$  に従うとき

$$Z = \frac{X - m}{\sigma}$$

とすると、確率変数  $Z$  は標準正規分布  $N(0, 1)$  に従います。

p.24 で、「コインを 100 回投げて表が出る回数を記録する」セット数を増やせば増やすほどきれいな山の形になり、そのきれいな形が正規分布であると言いました。

実は、 $X$  が二項分布  $B(n, p)$  に従うとき、 $n$  が十分大きければ正規分布に近づくことが知られています。  
 「十分大きい」ってどのくらい？という疑問をもつ人もいると思いますが、共通テストで出題される問題では、いつも「十分大きい」って考えて OK です。

つまり、 $X$  が二項分布  $B(n, p)$  に従い、かつ  $n$  が十分大きいとき、 $X$  は正規分布

$$N(np, np(1-p))$$

に近似的に従います。  
 二項分布の平均      二項分布の分散

$X$  が二項分布  $B(n, p)$  に従うとき

$$P(X=k) = {}_n C_k p^k (1-p)^{n-k} \quad (k=0, 1, 2, \dots, n)$$

となりました。でも、これを用いて確率を求めるのは結構骨が折れます。

試しに、 $X$  が二項分布  $B(300, \frac{1}{5})$  に従うとき、 $60 \leq X \leq 70$  となる確率  $P(60 \leq X \leq 70)$  を求めよう  
 とすると…

$$P(60 \leq X \leq 70) = \sum_{k=60}^{70} {}_{300} C_k \left(\frac{1}{5}\right)^k \left(1 - \frac{1}{5}\right)^{300-k} \quad \leftarrow \text{どう考えても大変ですよ。}$$

を計算せねばなりません。

しか～し！

二項分布を正規分布で近似できる場合には、正規分布表を使って確率を求められるので計算が簡単になるというわけです。

では、上の計算を実際にやってみましょう。

$X$  は二項分布  $B(300, \frac{1}{5})$  に従うから、 $X$  の平均  $E(X)$ 、分散  $V(X)$  は二項分布の公式より

$$E(X) = 300 \cdot \frac{1}{5} = 60, \quad V(X) = 300 \cdot \frac{1}{5} \left(1 - \frac{1}{5}\right) = 48$$

$B(300, \frac{1}{5})$  の 300 は十分に大きいので、 $X$  は正規分布  $N(60, 48)$  において求めればよいこととなります。 $X$  の平均は 60 で、 $X$  の標準偏差は  
 まよ、十分大きいって考えて OK です。

$$\sigma = \sqrt{V(X)} = \sqrt{48} = 4\sqrt{3}$$

だから

$$Z = \frac{X - E(X)}{\sigma} = \frac{X - 60}{4\sqrt{3}}$$

とおくと、 $Z$  は標準正規分布  $N(0, 1)$  に従います。

$60 \leq X \leq 70$  より

$$0 \leq X - 60 \leq 10$$

$$0 \leq \frac{X - 60}{4\sqrt{3}} \leq \frac{10}{4\sqrt{3}} = \frac{5\sqrt{3}}{6}$$

$$0 \leq Z \leq \frac{5 \cdot 1.732 \cdots}{6} \doteq 1.44$$

となり、 $P(60 \leq X \leq 70) = P(0 \leq Z \leq 1.44)$  だから、正規分布表より

$$P(0 \leq Z \leq 1.44) = 0.4251$$

であることがわかります。

つまり、 $P(60 \leq X \leq 70)$  のおよその値は 0.43 だということになります。  
 実際は約 0.456 なのですが、この程度の誤差は実用上問題ありません。

| $z_0$ | 0.00   | 0.01   | 0.02   | 0.03   | 0.04   |
|-------|--------|--------|--------|--------|--------|
| 0.0   | 0.0000 | 0.0040 | 0.0080 | 0.0120 | 0.0160 |
| 0.1   | 0.0398 | 0.0438 | 0.0478 | 0.0517 | 0.0557 |
| 0.2   | 0.0793 | 0.0832 | 0.0871 | 0.0910 | 0.0948 |
| 0.3   | 0.1179 | 0.1217 | 0.1255 | 0.1293 | 0.1331 |
| 0.4   | 0.1554 | 0.1591 | 0.1628 | 0.1664 | 0.1700 |
| 0.5   | 0.1915 | 0.1950 | 0.1985 | 0.2019 | 0.2054 |
| 0.6   | 0.2257 | 0.2291 | 0.2324 | 0.2357 | 0.2389 |
| 0.7   | 0.2580 | 0.2611 | 0.2642 | 0.2673 | 0.2704 |
| 0.8   | 0.2881 | 0.2910 | 0.2939 | 0.2967 | 0.2995 |
| 0.9   | 0.3159 | 0.3186 | 0.3212 | 0.3238 | 0.3264 |
| 1.0   | 0.3413 | 0.3438 | 0.3461 | 0.3485 | 0.3508 |
| 1.1   | 0.3643 | 0.3665 | 0.3686 | 0.3708 | 0.3729 |
| 1.2   | 0.3849 | 0.3869 | 0.3888 | 0.3907 | 0.3925 |
| 1.3   | 0.4032 | 0.4049 | 0.4066 | 0.4082 | 0.4099 |
| 1.4   | 0.4182 | 0.4207 | 0.4222 | 0.4236 | 0.4251 |
| 1.5   | 0.4332 | 0.4345 | 0.4357 | 0.4370 | 0.4382 |