3　正規分布

3.1　密度関数の特徴付け

密度関数をもつ絶対連続型データ／確率変数 X では、１つの数値単独での『相対度数（確率)』は０になっている事を見ました。すると次の様になりますから：

$$P[X \leq a] = P[X < a] + P[X = a] = P[X < a]$$

確率計算で不等号に等号がついている／いないは、密度関数をもつ場合に限っては気にする必要がありません。また

$$P[a \leq X \leq b] = P[X \leq b] - P[X < a] = P[X \leq b] - P[X \leq a]$$

とも書けますから、２パラメーターの確率 $P[a \leq X \leq b]$ は、１パラメーターの $P[X \leq c]$ 型の確率で表現出来る事が分かります。つまり、後者が任意の c に対して分かってしまえば、前者も分かると云う事です。従ってある関数 $h(x)$ が、任意の c に対して

$$P[X \leq c] = \int_{-\infty}^{c} h(x)dx \tag{3.1}$$

を満たしていさえいれば、任意の $a \leq b$ に対して

$$\begin{aligned} P[a \leq X \leq b] &= P[X \leq b] - P[X \leq a] \\ &= \int_{-\infty}^{b} h(x)dx - \int_{-\infty}^{a} h(x)dx \\ &= \int_{a}^{b} h(x)dx \end{aligned}$$

が得られ、$h(x)$ は X の密度関数である事が分かります。従って任意の c に対して (3.1) が成り立つ事を密度関数の定義としても問題ありません。１パラメーターの方がすっきりする事も多いでしょう（ただし広義積分になる可能性はあります）。

3.2　正規分布

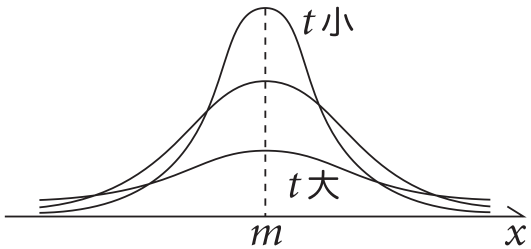
有名な Gauß積分：

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-y^2} dy = 1$$

において $y = \frac{x-m}{\sqrt{2t}}$ と変換すれば（ただし $t > 0$ とします）

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

が分かりますが、この被積分関数は非負であり、かつ数直線全体での積分が丁度１になっているので、密度関数の資格を満たしています。分布は m を中心にして左右対称であり、t が小さい時は中心付近で“尖った”高い山になり、t が大きい時は低くなだらかな山になっています。



定義 3.2.1 任意の実数 m と任意の正の実数 t に対して、密度関数が

$$\frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{(x-m)^2}{2t}}$$

である様な確率変数は平均 m、分散 t の正規分布（Normal distribution あるいは Gaussian distribution）に従っていると言い、この分布を記号 $N(m, t)$ で表します。ただしこの記号は、この講義においては『正規分布している確率変数』の意味で用いる事もあります。特に平均 0、分散 1 のものは標準正規分布と呼ばれます。

3.2.1　標準化

$N(m,t)$ に従う確率変数に定数 w を加えて得られる確率変数 $N(m,t)+w$ は、

$$\begin{aligned}P[N(m,t)+w \le u] &= P[N(m,t) \le u-w] \\ &= \int_{-\infty}^{u-w} \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{(x-m)^2}{2t}} dx \\ &= \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{\{(y-w)-m\}^2}{2t}} dy \\ &= \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{\{y-(m+w)\}^2}{2t}} dy\end{aligned}$$

が示す様に正規分布 $N(m+w,t)$ に従います。また、定数倍 $aN(m,t)$（$a>0$）も、

$$P[aN(m,t) \le u] = P\left[N(m,t) \le \frac{u}{a}\right] = \int_{-\infty}^{\frac{u}{a}} \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{(x-m)^2}{2t}} dx$$

において $x = \frac{y}{a}$ と置換すれば

$$P[aN(m,t) \le u] = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{\left(\frac{y}{a}-m\right)^2}{2t}} \frac{1}{a} dy = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi a^2 t}} \mathrm{e}^{-\frac{(y-am)^2}{2a^2 t}} dy$$

ですからこれは $aN(m,t)$ がやはり正規分布 $N(am, a^2 t)$ に従う事を示しています（$a<0$ の場合も対称性などを使って同様に計算出来るでしょう）。

以上から、正規分布の標準化はまた正規分布であり、確率変数 X が $N(m,t)$ に従うとき $\frac{X-m}{\sqrt{t}}$ は標準正規分布 $N(0,1)$ に従います。

3.2.2　標準正規分布表

正規分布に従う確率変数は様々な場面に現れてきますから、

$$P[v \le N(m,t) \le w] = \int_{v}^{w} \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{(x-m)^2}{2t}} dx$$

の右辺の積分値は頻繁に必要となります。しかしこの密度関数（＝被積分関数）の原始関数は初等関数ではないため、いわゆる普通の計算でこの積分値を求める事は出来ません。そこで、この積分値を（近似）計算した数表をあらかじめ用意しておくと便利です。

とは言っても、平均や分散が異なれば密度関数は異なりますから平均／分散ごとにいちいち数表を用意するかと言うとそんな馬鹿な事は致しません。標準化すれば、

$$P[v \le X \le w] = P\left[\frac{v-m}{\sqrt{t}} \le \frac{X-m}{\sqrt{t}} \le \frac{w-m}{\sqrt{t}}\right] = P\left[\frac{v-m}{\sqrt{t}} \le N(0,1) \le \frac{w-m}{\sqrt{t}}\right]$$

と変形出来ますから、標準正規分布に関する積分さえ分かっていれば一般の任意の正規分布に関する確率も計算出来るわけです。更には積分の性質と分布密度の対称性から

$$P[a \le N(0,1) \le b] = P[0 \le N(0,1) \le b] - P[0 \le N(0,1) \le a] \quad (b \ge a \ge 0)$$
$$P[0 \le N(0,1) \le a] = P[-a \le N(0,1) \le 0] \quad (a \ge 0)$$

でもありますから、結局このような

$$P[0 \le N(0,1) \le z] = \int_{0}^{z} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}} dx$$

の形の確率計算さえしておけば全ての場合に対応出来ます。これを数表にしたものがいわゆる標準正規分布表です。一番左の欄は z の値の小数点以下１桁までを表しており、一番上の欄が小数点以下２桁目を表しています。表中の確率は $P[N(0,1) \le z]$ ではなく $P[0 \le N(0,1) \le z]$ であることに注意して下さい。

問題 3.2.2　X が平均 4、分散 4 の正規分布に従うとき、標準正規分布表を参照して確率 $P[6 \le X \le 9]$ を求めて下さい。

$$\begin{aligned}P[6 \le X \le 9] &= P[6 \le N(4,4) \le 9] \\ &= P[2 \le N(0,4) \le 5] \\ &= P\left[1 \le N(0,1) \le \frac{5}{2}\right] \\ &= P[0 \le N(0,1) \le 2.5] - P[0 \le N(0,1) \le 1] \\ &\approx 0.4938 - 0.3413 \\ &= 0.1525\end{aligned}$$

□

問題 3.2.3 [高専教科書　例題 15.3] ある試験を 30000 人の受験者が受けました。100 点満点のところ平均点が 63.6、標準偏差が 13.4、点数の分布はほぼ正規分布でした。

(1) 80 点の受験者は約何番でしょうか。

(2) 10000 番の受験者の得点は約何点でしょうか。

---

(1) 得点を表す確率変数 X は正規分布 $N(63.6, 13.4^2)$ に従っていると仮定します。すると求めるべきものは 80 点より高得点の人が何人いるかであり、それは $P[80 < X]$ を求め、これに 30000 を掛ける事によって概算出来ます。これは

$$\begin{aligned} P[80 < X] &= P\left[\frac{80-63.6}{13.4} < \frac{X-63.6}{13.4}\right] \\ &= P[1.22 < N(0,1)] \\ &= 0.5 - P[0 \leq N(0,1) \leq 1.22] \end{aligned}$$

ですから正規分布表から

$$P[80 < X] = 0.5 - 0.3888 = 0.1112$$

が得られ、これに受験者総数の 30000 を掛けて 80 点より高得点の人が大体 3336 人だと分かります。従って 80 点の人は大体 3300 番程度であると考えられます。

(2) a 点以上の人が 10000 人居るとします。すると

$$\begin{aligned} \frac{10000}{30000} &= P[a \leq X] \\ 0.3333 &\approx P\left[\frac{a-63.6}{13.4} \leq \frac{X-63.6}{13.4}\right] \\ &\approx P\left[\frac{a-63.6}{13.4} \leq N(0,1)\right] \end{aligned}$$

ですが、a 点は平均点以上であると考えられますから

$$\begin{aligned} &\approx 0.5 - P\left[0 \leq N(0,1) \leq \frac{a-63.6}{13.4}\right] \\ 0.1667 &\approx P\left[0 \leq N(0,1) \leq \frac{a-63.6}{13.4}\right] \end{aligned}$$

となって正規分布表から $\frac{a-63.6}{13.4} \approx 0.43$、すなわち $a \approx 69.4$ が得られ、10000 番の受験者は大体 69 点であることが分かります。 □

注意 3.2.4 しかし母集団を正規分布と仮定してしまうと、得点は $-\infty$ から $+\infty$ までくまなく分布している事になってしまいます。

実際 $P[100 < X]$ を計算してみると

$$P[100 < X] = P\left[\frac{100-63.6}{13.4} < N(0,1)\right] \approx 0.0033$$

となっており、100 点を超えた人が 100 人も居る事になってしまいます。

また、$P[80 < X < 81]$ を計算してみると

$$P[80 < X < 81] = P\left[\frac{80-63.6}{13.4} < N(0,1) < \frac{81-63.6}{13.4}\right] \approx 0.0144$$

ですから『80 点を超えたが 81 点未満』と云う意味不明な人が 430 人も居るわけです。

離散データを絶対連続な（密度をもつ）分布で近似するとこのようなことが起きてしまいますが、むしろ元データの方を『正規分布を階級分けしたもの』と捉えた方が分かりやすいでしょうか。つまり、『80 点の人の相対度数』は 80 点を中央値とする幅 1 の階級の相対度数（確率）、$P[79.5 \leq X \leq 80.5]$ で計算されると考えるわけです。

そう考えれば

$$\begin{aligned} (\text{80 点の人の相対度数}) &= P[79.5 \leq X \leq 80.5] \\ &= P\left[\frac{79.5-63.6}{13.4} \leq N(0,1) \leq \frac{80.5-63.6}{13.4}\right] \\ &\approx P[1.19 \leq N(0,1) \leq 1.26] \\ &= P[0 \leq N(0,1) \leq 1.26] - P[0 \leq N(0,1) \leq 1.19] \\ &\approx 0.3962 - 0.3830 \\ &= 0.0132 \end{aligned}$$

となり、80 点の人数は 396 人と計算されます。

だから、(1) の計算も、

$$\begin{aligned} P[80.5 < X] &= P\left[\frac{80.5-63.6}{13.4} < \frac{X-63.6}{13.4}\right] \\ &= P[1.26 < N(0,1)] \end{aligned}$$

$$= 0.5 - P[0 \leq N(0,1) \leq 1.26]$$

ですから正規分布表から

$$P[80 < X] = 0.5 - 0.3962 = 0.1038$$

から 80 点より高得点の人の数は 3114 人程度であると考えるのが良さそうです。このような補正を半整数補正と呼んでいます。

(2) で半整数補正を行うと

$$\begin{aligned}
\frac{10000}{30000} &= P[a + 0.5 \leq X] \\
0.3333 &\approx P\left[\frac{a + 0.5 - 63.6}{13.4} \leq \frac{X - 63.6}{13.4}\right] \\
&\approx P\left[\frac{a + 0.5 - 63.6}{13.4} \leq N(0,1)\right] \\
&\approx 0.5 - P\left[0 \leq N(0,1) \leq \frac{a + 0.5 - 63.6}{13.4}\right] \\
0.1667 &\approx P\left[0 \leq N(0,1) \leq \frac{a + 0.5 - 63.6}{13.4}\right]
\end{aligned}$$

となって正規分布表から $\frac{a+0.5-63.6}{13.4} \approx 0.43$、すなわち $a \approx 68.9$ が得られ、10000 番の受験者は大体 69 点であることが分かります。

以下、半整数補正時の両側の誤差：

$$\begin{aligned}
P[X \leq -0.5] &= P\left[N(0,1) \leq \frac{-0.5 - 63.6}{13.4}\right] \\
&\approx P[N(0,1) \leq -4.78] \\
&= 0.5 - P[0 \leq N(0,1) \leq 4.78] \\
&\approx 0 \\
P[100.5 \leq X] &= P\left[\frac{100.5 - 63.6}{13.4} \leq N(0,1)\right] \\
&\approx P[2.75 \leq N(0,1)] \\
&= 0.5 - P[0 \leq N(0,1) \leq 2.75] \\
&= 0.5 - 0.4970 \\
&= 0.003
\end{aligned}$$

標準正規分布表

$$P[0 \leq N(0,1) \leq z] = \int_0^z \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}}\, dx$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.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
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998

3.3　問題演習

基本演習 3.1 [高専教科書　問題 15.3] X が平均 5、分散 4 の正規分布に従うとき、標準正規分布表を参照して次の確率を求めて下さい。
(2) $P[3 \leq X \leq 8]$　(3) $P[X \leq 10]$

基本演習 3.2 [高専教科書　問題 15.2] Z が標準正規分布に従うとき、標準正規分布表を参照して次の条件を満たす t の値を求めて下さい。
(2) $P[-t \leq Z \leq t] = 0.99$　(3) $P[Z \leq t] = 0.025$

基本演習 3.3 Z が標準正規分布に従うとき、標準正規分布表を参照することによって条件 $P[Z \geq t] = 0.015$ を満たす t の値を求めて下さい。

基本演習 3.4 ある試験を 30000 人の受験者が受けました。100 点満点のところ平均点が 63.6、標準偏差が 13.4、点数の分布はほぼ正規分布でした。
(1) 40 点の受験者はだいたい上から何番目ですか。
(2) 12000 番の受験者の得点はだいたい何点ですか。

発展演習 3.5 ある確率変数 X の密度関数が連続関数 $h(x)$ だったとします。このとき $h(x)$ は任意の x に対して $h(x) \geq 0$ を満たしており、また $\int_{-\infty}^{\infty} h(x)dx = 1$ でもある事を示して下さい。

発展演習 3.6 X が標準正規分布 $N(0,1)$ に従う時、$E[X^3] = 0$ であることを示して下さい。また、一般に $E[X^n]$ はどうなるかいろいろ計算してみて下さい。

発展演習 3.7 確率変数 X が標準正規分布 $N(0,1)$ に従うとき、X^2 はどんな密度関数をもつでしょうか。

発展演習 3.8 X の密度関数が $h(x)$ であるとき、$Y = 2X - 3$ の密度関数を $h(x)$ を使って表して下さい。

発展演習 3.9 X の密度関数が $h(x)$ であるとき、$Y = X^2$ の密度関数を $h(x)$ を使って表して下さい。

発展演習 3.10 Gauss 積分を使って、正規分布 $N(m,t)$ の平均値が m であり、分散が t である事を示して下さい。

問題演習解答

基本演習 3.1

(2)
$$\begin{aligned} P[3 \le X \le 8] &= P[-2 \le N(0,4) \le 3] \\ &= P\left[-1 \le \frac{N(0,4)}{2} \le \frac{3}{2}\right] \\ &= P[-1 \le N(0,1) \le 1.5] \\ &= P[0 \le N(0,1) \le 1.5] + P[0 \le N(0,1) \le 1] \\ &= 0.4332 + 0.3413 \\ &= 0.7745 \end{aligned}$$

(3)
$$\begin{aligned} P[X \le 10] &= P[X - 5 \le 5] \\ &= P\left[N(0,1) \le \frac{5}{2}\right] \\ &= 0.5 + P[0 \le N(0,1) \le 2.5] \\ &= 0.5 + 0.4938 \\ &= 0.9938 \end{aligned}$$

□

基本演習 3.2

(2)
$$\begin{aligned} 0.99 &= P[-t \le N(0,1) \le t] \\ &= 2P[0 \le N(0,1) \le t] \\ 0.495 &= P[0 \le N(0,1) \le t] \end{aligned}$$

なので標準正規分布表によれば、$t = 2.575$ です。

(3) まず与えられた式から $t < 0$ である事に注意します。すると

$$\begin{aligned} 0.025 &= P[N(0,1) \le t] \\ &= P[|t| \le N(0,1)] \\ 0.5 - 0.025 &= P[0 \le N(0,1) \le |t|] \\ 0.475 &= P[0 \le N(0,1) \le |t|] \end{aligned}$$

ですから標準正規分布表から $t = -1.96$ である事が分かります。 □

基本演習 3.3

確率の値が小さいので $t > 0$ であり、

$$\begin{aligned} 0.015 &= P[Z \ge t] \\ &= 0.5 - P[0 \le Z \le t] \\ 0.485 &= P[0 \le Z \le t] \end{aligned}$$

となりますから、正規分布表から $t \approx 2.17$ が分かります。 □

基本演習 3.4

(1) 点数の全体を X とすると、X は $N(63.6, 13.4^2)$ で近似されます。40 点以上の人の割合は

$$\begin{aligned} P[40 \le X] &\approx P[40 \le N(63.6, 13.4^2)] \\ &= P[-23.6 \le N(0, 13.4^2)] \\ &= P\left[-\frac{23.6}{13.4} \le N(0,1)\right] \\ &\approx P[-1.76 \le N(0,1)] \\ &= 0.5 + P[0 \le N(0,1) \le 1.76] \qquad \approx 0.9608 \end{aligned}$$

となりますから、その人数はおおよそ

$$30000 \times 0.9608 = 28824$$

となりますから、40 点だった人はおおよそ 28800 番であることが分かります。

(2) t 点以上の人が 12000 人いたとすると、

$$\begin{aligned} \frac{12000}{30000} &= P[t \le X] \\ 0.4 &\approx P[t \le N(63.6, 13.4^2)] \\ &= P\left[\frac{t - 63.6}{13.4} \le N(0,1)\right] \end{aligned}$$

ですが、確率が 0.5 より小さいので $t > 63.6$ であって、

$$0.4 \approx 0.5 - P\left[0 \le N(0,1) \le \frac{t - 63.6}{13.4}\right]$$

$$0.1 \approx P\left[0 \le N(0,1) \le \frac{t-63.6}{13.4}\right]$$

ですから、標準正規分布表から

$$\frac{t-63.6}{13.4} \approx 0.255 \quad \text{従って} \quad t \approx 67.017$$

が得られます。つまり、12000 番目の人はおおよそ 67 点であったと考えられます。 □

発展演習 3.5

連続関数がある点で負の値をとる場合、その近くでも負の値となっている事に注意。そうするとある区間の値をとる確率が負になってしまいます。

発展演習 3.6

奇数次は全て 0。偶数次は $E[X^{2n}] = (2n-1)!!$.

ただし !! は１つ飛ばしの階乗であり、例えば $9!! = 9 \cdot 7 \cdot 5 \cdot 3 \cdot 1$ など。便宜上 $0!! = (-1)!! = 1$ と定義する。

発展演習 3.7

X^2 は非負値しかとりませんから、密度関数は $x < 0$ では 0 になっているはずです。従って $t \ge 0$ に対して

$$P[X^2 \le t] = \int_0^t f(x)dx$$

となるような関数を見つければよく、

$$P[X^2 \le t] = P[-\sqrt{t} \le X \le \sqrt{t}] = \int_{-\sqrt{t}}^{\sqrt{t}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}} dx = 2\int_0^{\sqrt{t}} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}} dx$$

において $x^2 = z$ と置換すれば

$$P[X^2 \le t] = 2\int_0^t \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{z}{2}} \frac{1}{2\sqrt{z}} dz = \int_0^t \frac{1}{\sqrt{2\pi}} z^{-\frac{1}{2}} \mathrm{e}^{-\frac{z}{2}} dz$$

が得られますから、求める密度関数は以下の通りです：

$$\begin{cases} \frac{1}{\sqrt{2\pi}} x^{-\frac{1}{2}} \mathrm{e}^{-\frac{x}{2}} & 0 < x \\ 0 & x \le 0. \end{cases}$$

この分布は（自由度１の）カイ自乗分布と呼ばれています。 □

発展演習 3.8

X の密度関数が分かっていますので変形すれば

$$P[Y \le t] = P[2X - 3 \le t] = P\left[X \le \frac{t+3}{2}\right] = \int_{-\infty}^{\frac{t+3}{2}} h(x)dx = \int_{-\infty}^{t} h\left(\frac{y+3}{2}\right)\frac{1}{2}dy$$

となります。従って求める密度関数は $\frac{1}{2}h\left(\frac{x+3}{2}\right)$ です。 □

発展演習 3.9

発展演習 3.7 と同様な計算をすれば良いでしょう。

$$\begin{cases} \frac{h(\sqrt{x})+h(-\sqrt{x})}{2\sqrt{x}} & x > 0 \\ 0 & x \le 0 \end{cases}$$

□

発展演習 3.10

$$\begin{aligned} \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{(x-m)^2}{2t}} dx &= \int_{-\infty}^{\infty} (x-m) \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{(x-m)^2}{2t}} dx + \int_{-\infty}^{\infty} m \cdot \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{(x-m)^2}{2t}} dx \\ &= \left[-t \cdot \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{(x-m)^2}{2t}}\right]_{-\infty}^{\infty} + m \\ &= m \\ &= E[N(m,t)] \end{aligned}$$

$$\begin{aligned} &\int_{-\infty}^{\infty} (x-m)^2 \cdot \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{(x-m)^2}{2t}} dx \\ &= \int_{-\infty}^{\infty} \{-t(x-m)\} \cdot \left\{-\frac{x-m}{t}\right\} \cdot \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{(x-m)^2}{2t}} dx \\ &= \left[-t(x-m)\frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{(x-m)^2}{2t}}\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} (-t) \frac{1}{\sqrt{2\pi t}} \mathrm{e}^{-\frac{(x-m)^2}{2t}} dx \\ &= t \\ &= Var[N(m,t)] \end{aligned}$$

□