

確率分布のモーメント

確率変数 X の分布関数が $F(x)$ で表される時、期待値 $E(X)$ が

$$E(X) = \int_R x dF(x) \quad (1)$$

で表されることを説明しました。これは一般の可測関数 $\varphi(x)$ の分布関数 F に関するルベーグ積分

$$\int_R \varphi(x) dF(x) \quad (2)$$

で $\varphi(x)$ の特別な場合です。そこで、上記の積分を「関数 $\varphi(X)$ の期待値 $E(\varphi(X))$ 」と考えることにします。

関数 $\varphi(X)$ の期待値のうち、 $\varphi(X)$ が X の k 乗の形になっている特別な場合のものを、**確率変数 X の k 次モーメント (積率)** といいます。とくに、 $E(X^k)$ を**原点のまわりの k モーメント**とって μ'_k で表します。また、 X の期待値 (原点のまわりの 1 次モーメント) を μ で表すとき、 $E((X - \mu)^k)$ を**期待値のまわりの k 次モーメント**といい、 μ_k で表します。

モーメントは、その確率変数のしたがる確率分布の特徴、すなわち「どんなふうに分布しているか、どんなふうデータがばらばらか」を表現する量です。いちばん簡単なモーメントは、原点のまわりの 1 次モーメント μ'_1 で、これはすなわち上で述べた期待値です。また、期待値のまわりの 2 次モーメント μ_2 は、分布のばらつきの広さを表す量で、**分散**とよばれています¹。いつも期待値に近い値ばかりになる確率変数もあれば、期待値から遠く隔たった値にもしばしばなる確率変数もあります。このような「散らばり具合」を表すのが分散です。分散を $V(X)$ あるいは σ^2 という記号で表すこともあります。

さらに高次のモーメントも、分布のしかたを表すのに用いられます。その中でよく使われるのは、 $\alpha_3 = \mu_3/\sigma^3$ で定義される**歪度 (skewness)** と、 $\alpha_4 = \mu_4/\sigma^4$ を用いて $\alpha_4 - 3$ で定義される**尖度 (kurtosis)** です²。

$(X - \mu)^3$ は、 $X > \mu$ 、すなわち確率変数の値が期待値より大きいときは正で、 $X < \mu$ のときは負になります。したがって、確率変数の値が期待値より大きいときの密度関数の値が大きければ、 μ_3 は正になり、確率変数の値が期待値より小さいときの密度関数の値が大きければ μ_3 は負になりますから、歪度は、密度関数のグラフの、正負の方向への偏り具合をあらわします。

また、 $(X - \mu)^4$ は、確率変数の値が期待値に近いとき非常に小さくなりますから、単峰性分布 (密度関数のグラフの峰がひとつである分布) の場合に μ_4 の値が大きくなるためには、密度関数が $X = \mu$ 付近で突出して大きくなる必要があります。すなわち、尖度が大きいことは、密度関数のグラフが、 μ 付近で上にとがっていることを示しています。

¹分散の平方根を**標準偏差**といいます。

²”3” は、正規分布の場合の α^4 の値です。

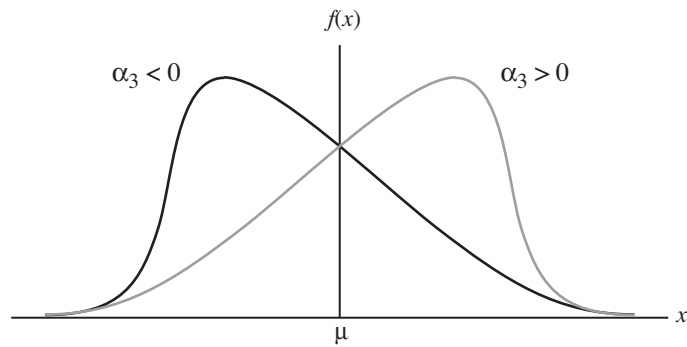


図 1: 密度関数と歪度

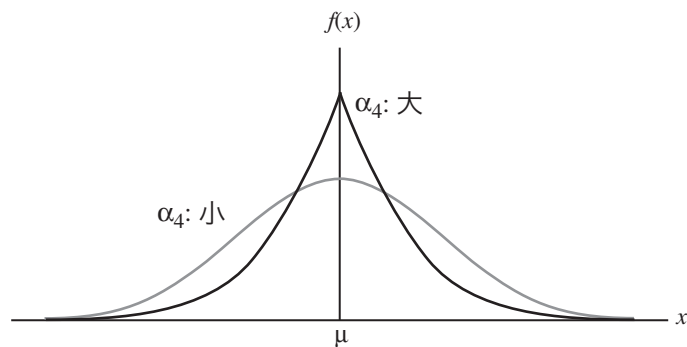


図 2: 密度関数と尖度

モーメント母関数

前節で示したモーメントの定義とは別に、1つの関数から各次のモーメントを生成する方法があります。この関数は**モーメント母関数**とよばれ、確率変数 X のモーメント母関数 $M_X(t)$ は

$$M_X(t) = E(e^{tX}) = \int_R e^{tx} dF(x) \quad (3)$$

で定義されます。ここで t は計算の便宜上導入された助変数です。

この関数からどうやってモーメントを導きだすかを見てみましょう。ただし、(3) 式の右辺は収束するとします。

e^{tx} をテイラー展開すると、(3) 式はつぎのように変形できます。

$$\begin{aligned} M_X(t) &= \int_R \left\{ 1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \cdots + \frac{(tx)^k}{k!} + \cdots \right\} dF(x) \\ &= \int_R 1 dF(x) + \int_R tx dF(x) + \int_R \frac{(tx)^2}{2!} dF(x) + \int_R \frac{(tx)^3}{3!} dF(x) + \cdots + \int_R \frac{(tx)^k}{k!} dF(x) + \cdots \\ &= \int_R dF(x) + t \int_R x dF(x) + \frac{t^2}{2!} \int_R x^2 dF(x) + \frac{t^3}{3!} \int_R x^3 dF(x) + \cdots + \frac{t^k}{k!} \int_R x^k dF(x) + \cdots \\ &= 1 + t\mu'_1 + \frac{t^2}{2!}\mu'_2 + \frac{t^3}{3!}\mu'_3 + \cdots + \frac{t^k}{k!}\mu'_k + \cdots \end{aligned} \quad (4)$$

(4) 式の最後の式を見ると、各項に原点の周りの各次数のモーメントが入っています。つまり、(3) 式の右辺が収束して、モーメント母関数が上のように t のべき級数に展開できるならば、モーメント母関数から各次数のモーメントが得られます。

実際に各次数のモーメントを求めるには、次のようにします。モーメント母関数を t で k 回微分すると、 t について $(k-1)$ 次以下の項は全て消えますから、

$$\frac{d^k}{dt^k} M_X(t) = \mu'_k + t\mu'_{k+1} + \frac{t^2}{2!}\mu'_{k+2} + \cdots + \frac{t^i}{i!}\mu'_{k+i} + \cdots \quad (5)$$

が得られます。ここで $t=0$ とおくと、

$$\frac{d^k}{dt^k} M_X(t)|_{t=0} = \mu'_k \quad (6)$$

となり、 k 次のモーメントが得られます。

ある確率分布が数式で与えられたときに、その期待値や分散を求める場合、モーメント母関数を求めてからそれらを導くほうが簡単になることがしばしばあります。今後の講義にも、いくつか例が出てくるとと思います。

特性関数

モーメント母関数 $M_X(t)$ に対して、 $i^2 = -1$ として

$$\phi_X(t) = M_X(it) = E(e^{itX}) \quad (7)$$

を**特性関数**といいます。特性関数の場合も、モーメント母関数と同様に、

$$\frac{d^k}{dt^k} \phi_X(t)|_{t=0} = i^k \mu'_k \quad (8)$$

とすることで、 k 次のモーメントを得ることができます。モーメント母関数の場合は、この関数の値が発散してしまうことがあります。特性関数の場合は $|e^{itX}| = 1$ ですから、必ず値が存在するという利点があります。

問題

- 「コインを 1 回投げて、表が出れば 20 円もらい裏が出れば 10 円払う」という賭けをします。
 - この賭け 1 回あたりにもらえる額は確率変数と考えられます。この確率変数を X で表すとき、 X がとりうる値をあげ、確率変数 X がしたがう確率分布を示してください。
 - この賭け 1 回あたりにもらえる額の期待値はいくらですか。
- 「コインを投げ続けて、はじめて表が出たらやめる。初めて表が出たのが 1 回目ならば 2^1 円、2 回目ならば 2^2 円、 \dots 、 n 回目ならば 2^n 円、 \dots の賞金がもらえる。」という賭けをします。
 - この賭け 1 回あたりにもらえる額を確率変数 X で表すとき、 X がとりうる値をあげ、 X がしたがう確率分布を示してください。
 - この賭け 1 回あたりの賞金の期待値はいくらですか。また、この賭けを 1 回やるのに 100 円払わなければならないとすると、あなたはこの賭けに参加しますか？