順序統計量

1. 順序統計量(指数関数)がよくわかる
2. 順序統計量の同時確率密度関数の期待値・分散がよくわかる
3. 順序統計量の中央値の確率密度関数がわかる
4. 順序統計量の幅の分布がわかる
5. 順序統計量の演習問題

順序統計量(指数関数)がよくわかる

【1】順序統計量のイメージが理解できる
順序統計量の基礎は、本冊子【順序統計量の考え方がよくわかる】で確認してください。

【2】順序統計量(指数関数)が理解できる

【問題】
確率変数Xは以下の指数分布に従うとする。その場合の順序統計量の期待値と分散値は以下となることを確認せよ。
$f(x)=e^{-x}$ (0 ≤ x < ∞)
$F(x)=1-e^{-x}$ (0 ≤ x < ∞)
(1)順序統計量の期待値　$E(X_i)=\sum_{r=1}^{i}\frac{1}{n-r+1}$
(2)順序統計量の分散　$V(X_i)=\sum_{r=1}^{i}\frac{1}{(n-r+1)^2}$

とても複雑な式になりますが、解いてみましょう。

★確率分布関数$f_{(i)}(x)$
確率分布関数$f_{(i)}(x)$=
$$\frac{n!}{(i-1)!1!(n-i)!}(1-e^{-x})^{i-1}[1-(1-e^{-x})]^{n-i}e^{-x}$$
$$=\frac{n!}{(i-1)!1!(n-i)!}(1-e^{-x})^{i-1}(e^{-x})^{n+1-i}$$
と素直に代入すればＯＫですね。

【3】順序統計量(指数関数)の期待値が計算できる
(1) 期待値$E(X_i)$
期待値$E(X_i)$は定義通り、
$$E(X_i)=\int_0^\infty xf_{(i)}(x)dx$$
$$=\int_0^\infty x\frac{n!}{(i-1)!1!(n-i)!}(1-e^{-x})^{i-1}(e^{-x})^{n+1-i}dx$$
$$=\frac{n!}{(i-1)!1!(n-i)!}\int_0^\infty x(1-e^{-x})^{i-1}(e^{-x})^{n+1-i}dx$$
=(式 1)

(式1)の$(1-e^{-x})^{i-1}$を、二項定理を使って展開します。
$(1-e^{-x})^{i-1}=\sum_{r=0}^{i-1}{}_{i-1}C_r1^r(-e^{i-1-r})$ですね。
($(p+q)^n=\sum_{r=0}^{n}{}_nC_rp^rq^{n-r}$と同じことをやっています。)

(式1)
$$=\frac{n!}{(i-1)!1!(n-i)!}\int_0^\infty x\sum_{r=0}^{i-1}{}_{i-1}C_r1^r(-e^{i-1-r})(e^{-x})^{n+1-i}dx$$
$$=\frac{n!}{(i-1)!1!(n-i)!}\int_0^\infty x\sum_{r=0}^{i-1}{}_{i-1}C_r(-1)^{i-1-r}(e^{-x})^{n-r}dx$$
=(式2)

(2) $\int_0^{\infty} xe^{-nx}dx$の計算

ここで、部分積分を実施します。⇒を微分する方向として、

$-\frac{1}{n}xe^{-nx} \Rightarrow xe^{-nx}$ -$\frac{1}{n}e^{-nx}$

$-\frac{1}{n^2}e^{-nx} \Rightarrow \frac{1}{n}e^{-nx}$

より、

$$\int_0^{\infty} xe^{-nx}dx = \left[-\frac{1}{n}xe^{-nx} - \frac{1}{n^2}e^{-nx}\right]_0^{\infty} = -\frac{1}{n^2}$$

となります。

n⇒n-r に変えて、(式 2)に代入します。

(式2)

$$= \frac{n!}{(i-1)!1!(n-i)!}\sum_{r=0}^{i-1} {}_{i-1}C_r(-1)^{i-1-r}\frac{1}{(n-r)^2}$$

よって、期待値$E(X_i)$は

$$E(X_i) = \frac{n!}{(i-1)!1!(n-i)!}\sum_{r=0}^{i-1} {}_{i-1}C_r(-1)^{i-1-r}\frac{1}{(n-r)^2}$$

となります。

ところで、問題を見ると、

$$E(X_i) = \sum_{r=1}^{i}\frac{1}{n-r+1}$$

と全く違う式です。

でも、代入すると値は一致します！不思議だけど

【$E(x_i)=\frac{n!}{(i-1)!(n-i)!}\sum_{r=0}^{i-1} {}_{i-1}C_r \quad (-1)^{i-1-r}\frac{1}{(n-r)^2}$】

n 5

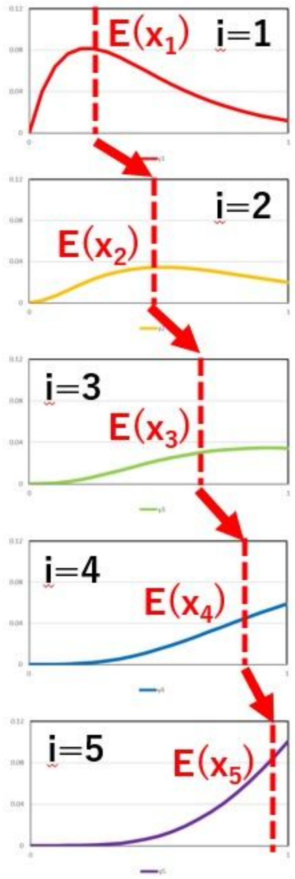
r/i	1	2	3	4	5
0	0.2	-0.8	1.2	-0.8	0.2
1		1.25	-3.75	3.75	-1.25
2			3.333333	-6.66667	3.333333
3				5	-5
4					5
合計	0.2	0.45	0.783333	1.283333	2.283333

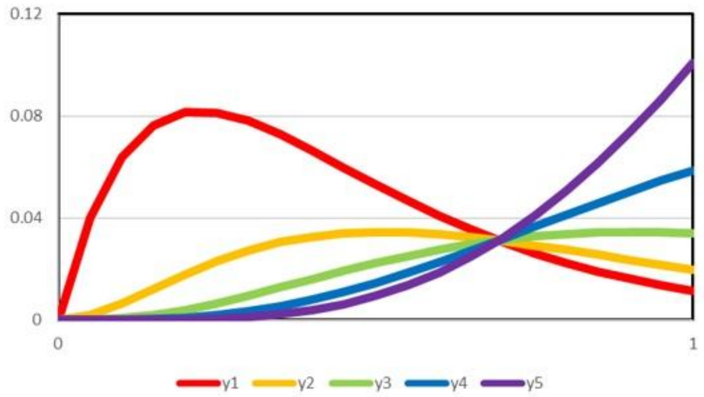
一致！

【$E(x_i)=\sum_{r=0}^{i}\frac{1}{n-r+1}$】

5	0.2	0.2	0.2	0.2	0.2
4		0.25	0.25	0.25	0.25
3			0.333333	0.333333	0.333333
2				0.5	0.5
1					1
合計	0.2	0.45	0.783333	1.283333	2.283333

(3) 期待値を可視化
図のようになります。





面白い事に、順序が増えることに　1/nずつ期待値が増えていきます。

$i=1$: $\frac{1}{5}$

$i=2$: $\frac{1}{5}+\frac{1}{4}$

$i=3$: $\frac{1}{5}+\frac{1}{4}+\frac{1}{3}$

$i=4$: $\frac{1}{5}+\frac{1}{4}+\frac{1}{3}+\frac{1}{2}$

…

【4】順序統計量(指数関数)の分散が計算できる

(1) 分散$V(X_i)$

まず期待値$E(X_i^2)$を求める必要がありますが、定義通り、

$$E(X_i^2)=\int_0^\infty x^2 f_{(i)}(x)dx$$

$$=\int_0^\infty x^2\frac{n!}{(i-1)!1!(n-i)!}(1-e^{-x})^{i-1}(e^{-x})^{n+1-i}dx$$

$$=\frac{n!}{(i-1)!1!(n-i)!}\int_0^\infty x^2(1-e^{-x})^{i-1}(e^{-x})^{n+1-i}dx$$

=(式3)

(式3)の$(1-e^{-x})^{i-1}$を、二項定理を使って展開します。
$(1-e^{-x})^{i-1}=\sum_{r=0}^{i-1}{}_{i-1}C_r 1^r(-e^{i-1-r})$ですね。
($(p+q)^n=\sum_{r=0}^{n}{}_nC_r p^r q^{n-r}$と同じことをやっています。)

(式3)
$$=\frac{n!}{(i-1)!1!(n-i)!}\int_0^\infty x^2\sum_{r=0}^{i-1}{}_{i-1}C_r 1^r(-e^{i-1-r})(e^{-x})^{n+1-i}dx$$
$$=\frac{n!}{(i-1)!1!(n-i)!}\int_0^\infty x^2\sum_{r=0}^{i-1}{}_{i-1}C_r(-1)^{i-1-r}(e^{-x})^{n-r}dx$$
=(式4)

(2) $\int_0^\infty x^2e^{-nx}dx$の計算
ここで、部分積分を実施します。⇒を微分する方向として、

$-\frac{1}{n}x^2e^{-nx}\Rightarrow x^2e^{-nx}-\frac{2}{n}xe^{-nx}$

$-\frac{2}{n^2}xe^{-nx}\Rightarrow\frac{2}{n}xe^{-nx}-\frac{2}{n^2}e^{-nx}$
$-\frac{2}{n^3}xe^{-nx}\Rightarrow\frac{2}{n^2}e^{-nx}$
より、
$$\int_0^\infty x^2e^{-nx}dx=\left[-\frac{1}{n}x^2e^{-nx}-\frac{2}{n^2}xe^{-nx}-\frac{2}{n^3}xe^{-nx}\right]_0^\infty=-\frac{2}{n^3}$$
となります。

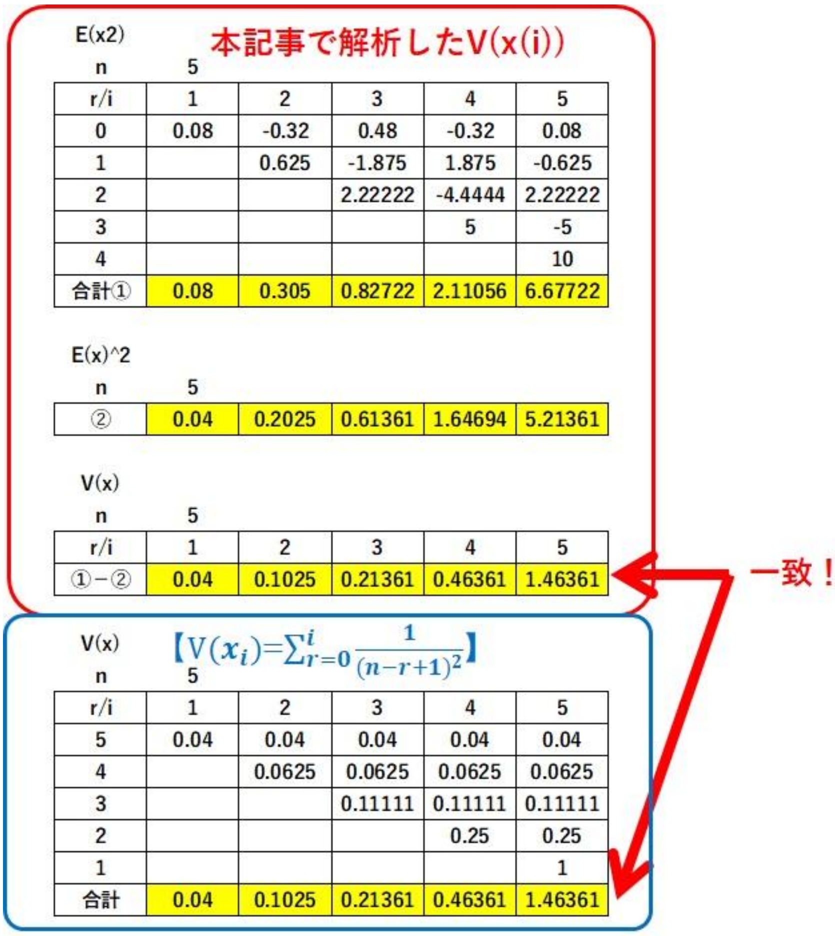
n⇒n-r に変えて、(式 4)に代入します。

(式4)
$$=\frac{n!}{(i-1)!(n-i)!}\sum_{r=0}^{i-1}{}_{i-1}C_r(-1)^{i-1-r}\frac{2}{(n-r)^3}$$
=(式5)

よって、分散$V(X_i)$は
(式5)-$E(X_i)^2$より、
$V(X_i)$
$$=\frac{n!}{(i-1)!(n-i)!}\sum_{r=0}^{i-1}{}_{i-1}C_r(-1)^{i-1-r}\frac{2}{(n-r)^3}-$$
$$\left(\frac{n!}{(i-1)!1!(n-i)!}\sum_{r=0}^{i-1}{}_{i-1}C_r(-1)^{i-1-r}\frac{1}{(n-r)^2}\right)^2$$
=(式6)
となります。
訳わからない式になりました。

ところで、問題を見ると、
$V(X_i)=\sum_{r=1}^{i}\frac{1}{(n-r+1)^2}$
と全く違う式です。

でも、n,iを代入すると値は一致します！不思議だけど。図の通りです。



本記事で解析したV(x(i))

E(x2)

n 5

r/i	1	2	3	4	5
0	0.08	-0.32	0.48	-0.32	0.08
1		0.625	-1.875	1.875	-0.625
2			2.22222	-4.4444	2.22222
3				5	-5
4					10
合計①	0.08	0.305	0.82722	2.11056	6.67722

E(x)^2

n 5

②	0.04	0.2025	0.61361	1.64694	5.21361

V(x)

n 5

r/i	1	2	3	4	5
①−②	0.04	0.1025	0.21361	0.46361	1.46361

一致！

V(x)

【$V(x_i)=\sum_{r=0}^{i}\frac{1}{(n-r+1)^2}$】

n 5

r/i	1	2	3	4	5
5	0.04	0.04	0.04	0.04	0.04
4		0.0625	0.0625	0.0625	0.0625
3			0.11111	0.11111	0.11111
2				0.25	0.25
1					1
合計	0.04	0.1025	0.21361	0.46361	1.46361

【5】自分で解いてわかった面白い事実

(1) 全く式が違うのに計算結果は同じとなったこと

●期待値$E(X_i)$は

$$E(X_i)=\frac{n!}{(i-1)!1!(n-i)!}\sum_{r=0}^{i-1} {}_{i-1}C_r(-1)^{i-1-r}\frac{1}{(n-r)^2}$$

と

$$E(X_i)=\sum_{r=1}^{i}\frac{1}{n-r+1}$$

は同じ値になります。

●分散$V(X_i)$は

●$V(X_i)$

$$=\frac{n!}{(i-1)!(n-i)!}\sum_{r=0}^{i-1} {}_{i-1}C_r(-1)^{i-1-r}\frac{2}{(n-r)^3}-\left(\frac{n!}{(i-1)!1!(n-i)!}\sum_{r=0}^{i-1} {}_{i-1}C_r(-1)^{i-1-r}\frac{1}{(n-r)^2}\right)^2$$

と、

$$V(X_i)=\sum_{r=1}^{i}\frac{1}{(n-r+1)^2}$$

が同じ結果になります。

数学的に一致する証明はこれからしますが、面白い結果が得られました。自分で実際解いてみるといろんなことが発見できますね。

以上、「順序統計量(指数関数)がよくわかる」を解説しました。

【1】順序統計量(一様分布)の同時確率密度関数の期待値・分散の導出
実例として一様分布で期待値・分散を計算します。他の分布関数ではちょっと計算が大変なので。
(1) 同時確率密度関数$f_{(i),(j)}(x_i,x_j)$
同時確率密度関数を確認しましょう。式を実際に使って慣れましょう。

$$f_{(i),(j)}(x_{(i)},x_{(j)})=C_{i,j}F(x_i)^{i-1}(F(x_j)-F(x_i))^{j-i-1}$$
$$(1-F(x_j))^{n-j}f(x_i)f(x_j)$$
$$(-\infty < x_i < x_j < \infty)$$

(2) 期待値・分散の導出例題
次の例題を考えます。

【例題】
確率変数Xの確率密度関数$f(x)$および分布関数$F(x)$が
●$f(x)$=1 (0 < x < 1)
●$F(x)$=x (0 < x < 1)
の一様分布に従うとする。このとき、$X_{(i)}$と$X_{(j)}$ (0 < $X_{(i)}$ < $X_{(j)}$ < 1)の同時分布について
(1)期待値E[$X_{(i)}$]
(2)期待値E[$X_{(i)}X_{(j)}$]
(3)分散V[$X_{(i)}$]
(4)共分散Cov[$X_{(i)}X_{(j)}$]
をそれぞれ求めよ。

【2】順序統計量(一様分布)の同時確率密度関数の期待値の導出
(1) 期待値の公式を確認
変数が 2 つx_i,x_jあるので、期待値は 3 種類考えます。

(i)E[$X_{(i)}$]
(ii)E[$X_{(j)}$]
(iii)E[$X_{(i)}X_{(j)}$]

(i)E[$X_{(i)}$]と(ii)E[$X_{(j)}$]は文字iとjを変えるだけで式は同じです。
なので、
(1)期待値E[$X_{(i)}$]⇒(i)E[$X_{(i)}$]
(2)期待値E[$X_{(i)}X_{(j)}$]⇒(iii)E[$X_{(i)}X_{(j)}$]

を解いていきます。

(2) 期待値 E[$X_{(i)}$]の導出

2変数ある同時確率密度関数は
$$f_{(i),(j)}(x_{(i)},x_{(j)})=C_{i,j}F(x_i)^{i-1}(F(x_j)-F(x_i))^{j-i-1}$$
$$(1-F(x_j))^{n-j}f(x_i)f(x_j)$$
ですが、１変数の確率密度関数は何でしょうか？

確率密度関数$f_{(i)}(x)=n_{n-1}C_{i-1}x^{i-1}(1-x)^{n-i}$ です。

期待値の計算結果だけ書くと

●期待値E$[X_{(i)}]$, E$[X_{(j)}]$
E$[X_{(i)}]$=$\frac{i}{n+1}$
E$[X_{(j)}]$=$\frac{j}{n+1}$
となります。スッキリした式ですね。

(3) 期待値 E$[X_{(i)}, X_{(j)}]$の導出
期待値 E$[x_i]$公式通りです。確認すると、

E$[x_{(i)}]$=$\int_0^1 x_{(i)} f_{(i)}(x)dx$

E$[x_{(i)}x_{(j)}]$=$\int_0^1 \int_0^{x_{(j)}} x_{(i)}x_{(j)} f_{(i),(j)}(x_{(i)}, x_{(j)})dx_{(i)}dx_{(j)}$

ですよね。
なお、
0 < $X_{(i)}$ < $X_{(j)}$ < 1
より、積分区間は
●$x_{(i)}$⇒0～$x_{(j)}$
●$x_{(j)}$⇒0～1
とします。

E$[x_{(i)}x_{(j)}]$=$\int_0^1 \int_0^{x_{(j)}} x_{(i)}x_{(j)} f_{(i),(j)}(x_{(i)}, x_{(j)})dx_{(i)}dx_{(j)}$

=$\int_0^1 \int_0^{x_{(j)}} x_{(i)}x_{(j)}$
$C_{i,j}x_{(i)}^{i-1}(x_{(j)} - x_{(i)})^{j-i-1}(1 - x_{(j)})^{n-j} \cdot 1 \cdot 1dx_{(i)}dx_{(j)}$
($F(x_i) = x_{(i)}$, $F(x_j) = x_{(j)}$, $f(x_i) = 1, f(x_j) = 1$を代入)

=
$\int_0^1 \int_0^{x_{(j)}} C_{i,j}x_{(i)}^{i}x_{(j)}(x_{(j)} - x_{(i)})^{j-i-1}(1 - x_{(j)})^{n-j}dx_{(i)}dx_{(j)}$
=(式1)

(式1)の$x_{(i)}$について先に積分します。つまり、
$\int_0^{x_{(j)}} x_{(i)}^{i}(x_{(j)} - x_{(i)})^{j-i-1}dx_{(i)}$
は第１種オイラーの積分
$\int_\alpha^\beta (x - \alpha)^m(\beta - x)^n dx$=$\frac{m!n!}{(m+n+1)!}(\beta - \alpha)^{m+n+1}$
を使って、
●x⇒$x_{(i)}$
●α⇒0
●m⇒i
●β⇒$x_{(j)}$
●n=$j - i - 1$
を代入します。

よって、(式1)の$x_{(i)}$についての積分部分は
$$\int_0^{x_{(j)}} x_{(i)}^i (x_{(j)} - x_{(i)})^{j-i-1} dx_{(i)}$$
$= \frac{i!(j-i-1)!}{j!} x_{(j)}^j$
=(式2)

(式2)を(式1)に代入します。
(式1)
$$= \int_0^1 C_{i,j} x_{(j)} (1-x_{(j)})^{n-j} \frac{i!(j-i-1)!}{j!} x_{(j)}^j dx_{(j)}$$
一旦係数を外に出します。
$$= C_{i,j} \frac{i!(j-i-1)!}{j!} \int_0^1 x_{(j)}^{j+1} (1-x_{(j)})^{n-j} dx_{(j)}$$
=(式3)

(式3)の積分を見ると、ベータ関数が使えることが分かります。
$$= \int_0^1 x_{(j)}^{j+1} (1-x_{(j)})^{n-j} dx_{(j)}$$
=B$(j+2, n-j+1)$
$= \frac{(j+1)!(n-j)!}{(n+2)!}$
=(式4)

(式4)を(式3)に代入し、$C_{i,j} = \frac{n!}{(i-1)!(j-i-1)!(n-j)!}$を代入すると
(式3)
$= \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \frac{i!(j-i-1)!}{j!} \frac{(j+1)!(n-j)!}{(n+2)!}$
$= \frac{i(j+1)}{(n+1)(n+2)}$

よって、

$$\text{期待値E}[X_{(i)} X_{(j)}] = \frac{i(j+1)}{(n+1)(n+2)}$$

計算できました。

期待値をまとめると

【期待値】
●期待値E$[X_{(i)}] = \frac{i}{n+1}$
●期待値E$[X_{(i)} X_{(j)}] = \frac{i(j+1)}{(n+1)(n+2)}$

【3】順序統計量(一様分布)の同時確率密度関数の分散の導出
(1) 分散の公式を確認
変数が2つx_i, x_jあるので、分散と共分散の計3種類考えます。
(i)V$[X_{(i)}]$
(ii)V$[X_{(j)}]$
(iii)Cov$[X_{(i)} X_{(j)}]$

(i)V[X_i]と(ii)V[X_j]は文字iとjを変えるだけで式は同じです。なので、
(1)分散V[$X_{(i)}$]⇒(i)E[X_i]
(2)分散Cov[$X_{(i)}X_{(j)}$]⇒(iii)Cov[$X_{(i)}X_{(j)}$]
を解いていきます。

(2) 分散$V[X_{(i)}]$の導出
分散の計算結果だけ書くと

●分散V[$X_{(i)}$],V[$X_{(j)}$]

$$V[X_{(i)}]=\frac{i(n-i+1)}{(n+1)^2(n+2)}$$

$$V[X_{(j)}]=\frac{j(n-j+1)}{(n+1)^2(n+2)}$$

となります。</div>

(3) 共分散$V[X_{(i)},X_{(j)}]$の導出
共分散V[X,Y]=E[XY]-E[X]E[Y]という公式があり、
V[$X_{(i)}X_{(j)}$]= E[$X_{(i)}X_{(j)}$ - E[$X_{(i)}$] E[$X_{(j)}$]
から計算できます。
よって、

$$V[X_{(i)}X_{(j)}]=E[X_{(i)}X_{(j)} - E[X_{(i)}]E[X_{(j)}]$$

$$=\frac{i(j+1)}{(n+1)(n+2)} - \frac{i}{n+1}\ \frac{j}{n+1}$$

$$=\frac{i(n-j+1)}{(n+1)^2(n+2)}$$

よくみると、

$$V[X_{(i)}]=\frac{i(n-i+1)}{(n+1)^2(n+2)}$$

$$V[X_{(i)}X_{(j)}]=\frac{i(n-j+1)}{(n+1)^2(n+2)}$$

分子を比較するとよく似ているのがわかります。

分散、共分散をまとめると

【分散、共分散】

●期待値V[$X_{(i)}$]= $\frac{i(n-i+1)}{(n+1)^2(n+2)}$

●期待値V[$X_{(i)}X_{(j)}$] =$\frac{i(n-j+1)}{(n+1)^2(n+2)}$

順序統計量(一様分布)の同時確率密度関数の難しい式を使う良い演習ができました。
以上、「順序統計量の同時確率密度関数の期待値・分散がよくわかる」を解説しました。

【1】順序統計量の中央値の確率密度関数の導出

(1) 順序統計量の中央値の導出

確率変数$X_1, X_2, \cdots, X_n$の中央値は、nが奇数のときと、偶数のときで表示が若干変わります。

(i)奇数のとき：$\frac{n+1}{2}$

(ii)偶数のとき：$\frac{\frac{n}{2}+(\frac{n}{2}+1)}{2}$

例として、n =7 のときは、中央は 4 です。 $\frac{7+1}{2}$=4 で計算できます。

n =8 のときは、中央は 4 と 5 の真ん中です。$\frac{\frac{8}{2}+\frac{8}{2}+1}{2}$=4.5 で計算できますね。

これを順序統計量の確率密度関数に代入すればＯＫです。

(2) 順序統計量の中央値の確率密度関数の導出

nが奇数の場合

$f_{(i)}(x)=\frac{n!}{(i-1)!1!(n-i)!}F(x)^{i-1}(1-F(x))^{n-i}f(x)$に代入します。

$$f_M(x)=\frac{n!}{(\frac{n+1}{2}-1)!1!(n-\frac{n+1}{2})!}F(x)^{\frac{n+1}{2}-1}(1-F(x))^{n-\frac{n+1}{2}}f(x)$$
$$=\frac{n!}{(\frac{n-1}{2})^2}(F(x)(1-F(x))^{\frac{n-1}{2}}f(x)$$

となります。

つまり、

$$f_M(x)=\frac{n!}{(\frac{n-1}{2})^2}(F(x)(1-F(x))^{\frac{n-1}{2}}f(x)$$

nが偶数の場合

$f_{(i)}(x)=\frac{n!}{(i-1)!1!(n-i)!}F(x)^{i-1}(1-F(x))^{n-i}f(x)$に代入します。

ここで、式の見た目をよくするために１つトリックを仕込みます。(良くないと思うが)

$(i-1)$のiに$i=\frac{n}{2}$を
$(n-i)$のiに$i=\frac{n}{2}+1$を
代入して、共に
$\frac{n}{2}-1$とさせます。

また、(f(x))の部分を(f(x_1)), (f(x_2))に分けます。

$$f_M(x)=\frac{n!}{(\frac{n}{2}-1)!1!(\frac{n}{2}-1)!}F(x_1)^{\frac{n}{2}-1}(1-F(x_2))^{\frac{n}{2}-1}f(x_1)f(x_2)$$
$$=\frac{n!}{((\frac{n}{2}-1)!}^{2}(F(x_1)(1-F(x_2))^{\frac{n}{2}-1}f(x_1)f(x_2)$$

となります。

つまり、

$$f_M(x)=\frac{n!}{((\frac{n}{2}-1)!}^{2}(F(x_1)(1-F(x_2))^{\frac{n}{2}-1}f(x_1)f(x_2)$$

まとめると、

(i)奇数のとき：$f_M(x) = \frac{n!}{(\frac{n-1}{2})^2}(F(x)(1-F(x))^{\frac{n-1}{2}}f(x)$

(ii)偶数のとき：$f_M(x) = \frac{n!}{((\frac{n}{2}-1)!}^2$

$(F(x_1)(1-F(x_2))^{\frac{n}{2}-1}f(x_1)f(x_2)$

【2】順序統計量の中央値の確率密度関数の例題

一様分布を使って、具体的な中央値の確率密度関数の式を作ってみましょう。

【問】

確率変数Xの確率密度関数$f(x)$および分布関数$F(x)$は

●$f(x)$=1 (0 < x < 1)

●$F(x)$=x (0 < x < 1)

である一様分布に従うとする。

(1)　中央値の確率密度関数を導出せよ。

(2) 期待値と分散を求めよ。(nが奇数の場合のみでよい)

(1)順序統計量の中央値の確率密度関数の導出

(i)奇数のとき：$f_M(x) = \frac{n!}{(\frac{n-1}{2})!^2}(F(x)(1-F(x))^{\frac{n-1}{2}}f(x)$

(ii)偶数のとき：$f_M(x) = \frac{n!}{((\frac{n}{2}-1)!^2}$

$(F(x_1)(1-F(x_2))^{\frac{n}{2}-1}f(x_1)f(x_2)$

に$f(x)=1$,$F(x)=x$を代入します。

nが奇数の場合

$f_M(x) = \frac{n!}{(\frac{n-1}{2}!)^2}(x(1-x)^{\frac{n-1}{2}}$

(0 < x < 1)

nが偶数の場合

$f_M(x) = \frac{n!}{((\frac{n}{2}-1)!)^2}(x_1(1-x_2)^{\frac{n}{2}-1}$

(0 < x < 1)

(2)期待値と分散の導出

期待値と分散は

●期待値E=$\frac{i}{n+1}$

●分散V=$\frac{i(n-i+1}{(n+1)^2(n+2)}$

この式に$i = \frac{n+1}{2}$を代入して、

●期待値E=$\frac{i}{n+1} = \frac{\frac{n+1}{2}}{n+1} = \frac{1}{2}$

●分散V=$\frac{i(n-i+1)}{(n+1)^2(n+2)}$

$= \frac{1}{4(n+2)}$

以上、「順序統計量の中央値の確率密度関数がわかる」を解説しました。

順序統計量の幅の分布がわかる

【1】順序統計量の幅の分布の確率密度関数がわかる</h2>
(1) R 管理図の係数導出に必須
順序統計量から「R 管理図の係数導出」を求めますが、その導出過程が難しいです。

(2) 同時確率密度関数からスタート
基本は同時確率密度関数からスタートします。基本的な内容は関連記事でご確認ください。

【関連記事】順序統計量の同時確率密度関数の導出がよくわかる https://qcplanets.com/method/statistics-method/order-statistics-base5/

この記事では、同時確率密度関数の式からスタートします。

$$f_{(i),(j)}(x_{(i)},x_{(j)})=C_{i,j}F(x_i)^{i-1}(F(x_j)-F(x_i))^{j-i-1}$$
$$(1-F(x_j))^{n-j}f(x_i)f(x_j)$$

(3) 順序統計量の幅の分布の確率密度関数
同時確率分布の確率密度関数において、

$$W=X_{(n)}-X_{(1)}$$

と定義して、幅 W の分布を考えます。
まず、

$$f_{(i),(j)}(x_{(i)},x_{(j)})=C_{i,j}F(x_i)^{i-1}(F(x_j)-F(x_i))^{j-i-1}$$
$$(1-F(x_j))^{n-j}f(x_i)f(x_j)$$

へ、i=1、j=nを代入します。

$$f_{(1),(n)}(x_{(1)},x_{(n)})=C_{1,n}(F(x_n)-F(x_1))^{n-2}f(x_1)f(x_n)$$

まとめると、$x_n=x_1+w$に注意して

$$f_W(w)=n(n-1)(F(x_1+w)-F(x_1))^{n-2}f(x_1)f(x_1+w)$$

=(式 1)
となります。この式を使います。

(4) 順序統計量の幅の分布の確率密度関数を積分

$$f_W(w)=n(n-1)(F(x_1+w)-F(x_1))^{n-2}f(x_1)f(x_1+w)$$

はまだ、変数x_1がいて、変数wだけの式ではないので、変数x_1で積分します。

Wについての確率密度関数は

$$f_W(w)=n(n-1)\int_{-\infty}^{\infty}(F(x_1+w)-F(x_1))^{n-2}$$
$$f(x_1)f(x_1+w)dx_1$$

=(式2)
となります。
さらに(式2)を積分すると、変数Wについての確率も導出できます。

$$\Pr(W)=n\int_{-\infty}^{\infty}(F(x_1+w)-F(x_1))^{n-1}f(x_1)dx_1$$

=(式3)
となります。

(式2)⇒(式3)においては、
$f(x_1+w)$= $\frac{d(F(x_1+w)-F(x_1))}{dx}$として、
$(n-1)(F(x_1+w)-F(x_1))^{n-2}f(x_1+w)$を
$(n-1)(F(x_1+w)-F(x_1))^{n-2}\frac{d(F(x_1+w)-F(x_1))}{dx}$として、
$(F(x_1+w)-F(x_1))$を$(n-2)$乗から$(n-1)$乗に１つ指数を上げる計算をしています。

後で使う式なのでまとめると、

●(式1)：順序統計量の幅の分布の確率密度関数
$f_W(w)$=$n(n-1)(F(x_1+w)-F(x_1))^{n-2}f(x_1)f(x_1+w)$
●(式2)：順序統計量の幅の分布の確率密度関数
$f_W(w)$= $n(n-1)\int_{-\infty}^{\infty}(F(x_1+w)-F(x_1))^{n-2}$
$f(x_1)f(x_1+w)dx_1$
●(式3)：順序統計量の幅の分布の確率
Pr(W)=$n\int_{-\infty}^{\infty}(F(x_1+w)-F(x_1))^{n-1}f(x_1)dx_1$

【2】順序統計量の幅の分布の例題
実際に例題を見ながら計算してみましょう。
(1) 順序統計量の幅の分布の例題

【例題】
確率変数Xの確率密度関数$f(x)$および分布関数$F(x)$は
●$f(x)=1$ (0 < x < 1)
●$F(x)=x$ (0 < x < 1)
である一様分布に従うとする。この時の、
(i)W=$X_{(n)}$- $X_{(1)}$と定義する変数Wの確率密度関数$f_W(w)$を求めよ。
(ii) W=$X_{(j)}$- $X_{(i)}$と定義する変数W同時確率密度関数
$f_W(i,j)(w)$
をそれぞれ求めよ。

公式から実際に解いて、イメージをつけましょう。
(i)確率密度関数$f_W(w)$の解法
(式2)に$F(x)$=x,$f(x)$=1を代入します。

(式2)
=$f_W(w)$= $n(n-1)\int_{-\infty}^{\infty}(F(x_1+w)-F(x_1))^{n-2}$
$f(x_1)f(x_1+w)dx_1$
=$n(n-1)\int_{0}^{1-w}((x_1+w)-x_1)^{n-2}dx_1$
=$n(n-1)w^{n-2}(1-w)$ (0 < w < 1)

wの式ですが、x_1で積分する点に注意しましょう。
また、$x_n=x_1+w$ < 1より x_1 < 1-wで積分します。

(ii)同時確率密度関数$f_W(i,j)(w)$の解法
同時確率密度関数は

同時確率密度関数は
$f_{(i),(j)}(x_{(i)}, x_{(j)})$=$C_{i,j} F(x_i)^{i-1} (F(x_j) - F(x_i))^{j-i-1}$
$(1-F(x_j))^{n-j} f(x_i) f(x_j)$
より、$f(x)=1$、$F(x)=x$を代入すると、

$f_{W(i,j)}(w)$=$C_{i,j} F(x_i)^{i-1} (F(x_j) - F(x_i))^{j-i-1}$
$(1-F(x_j))^{n-j} f(x_i) f(x_j)$
$x_j = x_i + w$と置くと、

$f_{W(i,j)}(w)$=$C_{i,j} F(x_i)^{i-1} (F(x_i + w) - F(x_i))^{j-i-1}$
$(1-F(x_i + w))^{n-j} f(x_i) f(x_i + w)$

さらに、$F(x) = x, f(x) = 1$を代入すると、

$f_{W(i,j)}(w)$=$C_{i,j} x_i^{i-1} (w)^{j-i-1} (1 - x_i - w))^{n-j}$

この式をdx_iで積分すると、
$x_n = x_1 + w$ < 1より x_1 < 1-wで積分する点に注意して、

$f_{W(i,j)}(w)$=$C_{i,j} \int_0^{1-w} x_i^{i-1} (w)^{j-i-1} (1 - x_i - w))^{n-j} dx_i$
=$C_{i,j} w^{j-i-1} (1-w)^{n-j+i}$
となります。

となります。

幅の分布で具体的な関数になっても、いまいちピンと来ませんが、計算はできますね。

以上、「順序統計量の幅の分布がわかる」を解説しました。

順序統計量の演習問題

【1】順序統計量と確率変数の変換の演習問題

(1) 演習問題

3つの確率変数Y_1、Y_2、Y_3について、
$Y_1 < Y_2 < Y_3$
が成り立ち、互いに独立で、一様分布[0,1]に従うとする。この場合、
$Z=Y3-Y1$
の確率密度関数を求めよ。

(2) 解法のポイント

ポイントは、２つあり、
(i) 確率変数の変換(2 変数)
(ii) 順序統計量の同時確率密度関数
を使います。

(3) 解法

一様分布からY_iの確率密度関数$f_i(y)$と分布関数$F_i(y)$は
●$f_i(y)$=1 (0 < y < 1)
●$F_i(y)$=y (0 < y < 1)

です。

次にy_1とy_3の同時確率密度関数は公式
$f_{(i),(j)}(x_{(i)},x_{(j)})$=$C_{i,j}F(x_i)^{i-1}(F(x_j)-F(x_i))^{j-i-1}$
$(1-F(x_j))^{n-j}f(x_i)f(x_j)$
から
$f_{(1),(3)}(y_{(1)},y_{(3)})$=$C_{1,3}F(y_1)^{1-1}(F(y_3)-F(y_1))^{3-1-1}$
$(1-F(x_3))^{3-3}f(x_1)f(x_3)$
$f_{(1),(3)}(y_{(1)},y_{(3)})$=6$(F(y_3)-F(y_1))$
=$6(y_3-y_1)$

そして、$z=y_3-y_1$より、２変数の変換をします。
●$z=y_3-y_1$
●w=$y1$
とおいて、変換すると、
$f_{z,w}(z,w)$= $f(y_1(z,w),y_3(z,w))|detJ|$で
=$6(y_3-y_1)\cdot 1$
=$6z$ (ただし、0 < z < w < 1)

ここでヤコビアン行列から
Jは

$$J=\begin{pmatrix} \frac{\partial y_1}{\partial z} & \frac{\partial y_3}{\partial w} \\ \frac{\partial y_1}{\partial z} & \frac{\partial y_3}{\partial w} \end{pmatrix}$$

ヤコビアン行列を実際に代入すると、

$$J=\begin{pmatrix}-1 & 1\\ 1 & 0\end{pmatrix}$$

から行列式は-1 で、その絶対値は 1 です。

最後に、z の周辺確率密度関数を計算します。
$f_{z,w}(z,w)$=$6z$ (ただし、0 < z < w < 1)
は変数z,wの式なので、無関係なwについて積分します。

積分区間は、z < w < 1に注意します。

$f(z)$=$\int_z^1 6zdz$
=$6z(1-z)$
(答え)
(答え)
できましたね！

【2】順序統計量の具体例の演習問題
(1) 演習問題

> 3つの確率変数x_1、x_2、x_3について、
> x_1 < x_2 < x_3
> が成り立ち、互いに独立で、以下の確率密度関数$f_i(x)$
> ●$f_i(x)=2x$ (0 < x < 1)
> に従う。
> (1) x_1、x_2、x_3それぞれの確率密度関数を導出せよ。
> (2) x_1、x_2、x_3それぞれの期待値E[x_i]を計算し、
> E[x_1] < E[x_2] < E[x_3]を確認せよ。

1 変数の順序統計量の確率密度関数と期待値の計算ですね。関連記事でも解説しています。

(2) 解法
まず、分布関数$F_i(x)$は確率密度関数$f_i(x)$を積分して、
●$f_i(x)=2x$ (0 < x < 1)
●$F_i(x)=x^2$ (0 < x < 1)
ですね。
①確率密度関数$f_i(x)$を計算
確率密度関数$f_i(x)$を計算します。
$f_i(x)$=$\frac{n!}{(i-1)!1!(n-i)!}F(x)^{i-1}[1-F(x)]^{n-i}f(x)$
より、

●i=1のとき、
$f_1(x)$=$\frac{3!}{(1-1)!1!(3-1)!}F_1(x)^{1-1}[1-F_1(x)]^{3-1}f_1(x)$
=$\frac{3!}{2!}[1-F_1(x)]^2 f_1(x)$
=$3(1-x^2)^2\cdot 2x$
=$6x(1-x^2)^2$

●i=2のとき、

$f_2(x)=\frac{3!}{(2-1)!1!(3-2)!}F_2(x)^{2-1}[1-F_2(x)]^{3-2}f_2(x)$

$=3!F_2(x)[1-F_2(x)]f_2(x)$

$=6x^2(1-x^2)\cdot 2x$

$=12x^3(1-x^2)$

●i=3のとき、

$f_3(x)=\frac{3!}{(3-1)!1!(3-3)!}F_3(x)^{3-1}[1-F_3(x)]^{3-3}f_3(x)$

$=\frac{3!}{2!}F_3(x)^2 f_3(x)$

$=3(x^2)^2\cdot 2x$

$=6x^5$

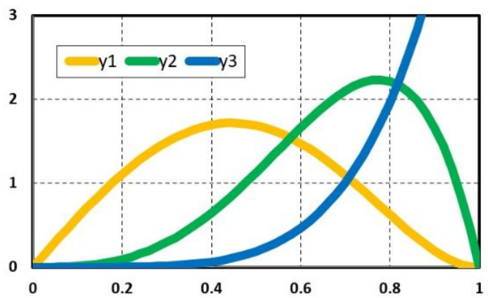
3つの関数をまとめると

●$f_1(x)=6x(1-x^2)^2$

●$f_2(x)=12x^3(1-x^2)$

●$f_3(x)=6x^5$

図で描くと、E[x_1] < E[x_2] < E[x_3]に見えますよね。



②期待値 E[x_i]を計算

$$\mathrm{E}[x_1]=\int_0^1 xf_1(x)dx$$
$$=\int_0^1 6x^2(1-x^2)^2dx$$
$$=\frac{16}{35}$$

$$\mathrm{E}[x_2]=\int_0^1 xf_2(x)dx$$
$$=\int_0^1 12x^4(1-x^2)dx$$
$$=\frac{24}{35}$$

$$\mathrm{E}[x_3]=\int_0^1 xf_3(x)dx$$
$$=\int_0^1 6x^6dx$$
$$=\frac{30}{35}$$

$\frac{16}{35}<\frac{24}{35}<\frac{30}{35}$ より

E[x_1] < E[x_2] < E[x_3]が確認できた。

【3】順序統計量の範囲 R の演習問題(Ⅰ)

(1) 演習問題

n個の確率変数X_1、X_2、…、X_nが定数λの指数分布に従い、
X_1、X_2、…、X_nが順序統計量に従うとする。
(1) X_1の確率密度関数を求めよ。
(2) X_nの確率密度関数を求めよ。
(3) $R = X_n - X_1$の確率密度関数を求めよ。

(2)解法のポイント

ポイントは、
①順序統計量の同時確率密度関数
②順序統計量の幅の分布
を使います。

(3)解法

①確率密度関数と分布関数

それぞれ先に導出しておくと、
●$f(x)=\lambda e^{-\lambda x}$ $(x>0)$
●$F(x)=1-e^{-\lambda x}$ $(x>0)$
ですね。

② (1)の解法

公式
$f_{(i)}(x)=\frac{n!}{(i-1)!1!(n-i)!}F(x)^{i-1}[1-F(x)]^{n-i}f(x)$から

$f_{(1)}(x)=\frac{n!}{(1-1)!1!(n-1)!}F(x)^{1-1}[1-F(x)]^{n-1}f(x)$
$=\frac{n!}{(n-1)!}[1-F(x)]^{n-1}f(x)$
$=n[1-(1-e^{-\lambda x})]^{n-1}\cdot\lambda e^{-\lambda x}$
$=n\lambda e^{-n\lambda x}$

② (2)の解法

公式
$f_{(i)}(x)=\frac{n!}{(i-1)!1!(n-i)!}F(x)^{i-1}[1-F(x)]^{n-i}f(x)$から

$f_{(n)}(x)=\frac{n!}{(n-1)!1!(n-n)!}F(x)^{n-1}[1-F(x)]^{n-n}f(x)$
$=\frac{n!}{(n-1)!}F(x)^{n-1}f(x)$
$=n^{n-1}f(x)$
$=n(1-e^{-\lambda x})^{n-1}\lambda e^{-\lambda x}$
$=n\lambda(1-e^{-\lambda x})^{n-1}e^{-\lambda x}$

③(3)の解法

$R=X_n-X_1$より幅の分布を考えます。

Rは、同時確率密度関数を用いて、
公式
$f_{(i),(j)}(x_{(i)},x_{(j)})$=$C_{i,j}F(x_i)^{i-1}(F(x_j)-F(x_i))^{j-i-1}$
$(1-F(x_j))^{n-j}f(x_i)f(x_j)$
を使って

$f_R(r)$= $f_{(1),(n)}(x_1,x_n)$
=$n(n-1)[(F(x_1+r)-F(x_1)]^{n-2}f(x_1)f(x_1+r)$
=
$n(n-1)[(1-e^{-\lambda(x_1+r)})-(1-e^{-\lambda x_1})]^{n-2}\lambda e^{-\lambda x_1}\lambda e^{-\lambda(x_1+r)}$
=$n(n-1)\lambda^2 e^{-\lambda(2x_1+r)}(e^{-\lambda x_1}-e^{-\lambda(x_1+r})^{n-2}$
=(式1)

(式1)をrについて積分し、x_1をxに書き直します。

Rの確率密度関数は
$f_R(x)$= $\int_0^\infty$ (式$1)dr$
= $\int_0^\infty n(n-1)\lambda^2 e^{-\lambda(2x_1+r)}(e^{-\lambda x_1}-e^{-\lambda(x_1+r})^{n-2}dr$
=$n(n-1)\lambda^2 e^{-n\lambda x}\int_0^\infty e^{-\lambda r}(1-e^{-\lambda r})^{n-2}dr$
=(式2)

ここで、$e^{-\lambda r}=t$とおくと、dt=$-\lambda tdr$となり、(式2)に代入します。

(式2)
=$n(n-1)\lambda^2 e^{-n\lambda x}-\frac{1}{\lambda}\int_1^0(1-t)^{n-2}dt$
=$n(n-1)\lambda e^{-n\lambda x}\int_0^1(1-t)^{n-2}dt$
=$n\lambda e^{-n\lambda x}$
となります。

できましたね！

【4】順序統計量の範囲 R の演習問題(Ⅱ)

(1) 演習問題</h3>

確率変数X_1,X_2,...X_nを[0,1]の一様分布に従う順序統計量とする。
標本範囲R=X_n-X_1を考える。
(1) 標本範囲R確率密度関数$f_R(x)$を求めよ。
(2) 標本範囲Rの期待値E[R]を計算せよ。
(3) 標本範囲Rが半分の$\frac{1}{2}$より大きくなる確率を求めよ。

(2)解法

① (1)の解法

(式2)に$F(x)$=x,$f(x)$=1を代入します。

(式2)

=$f_W(w)$= $n(n-1)\int_{-\infty}^{\infty}(F(x_1+w)-F(x_1))^{n-2}$

$f(x_1)f(x_1+w)dx_1$

=$n(n-1)\int_{0}^{1-w}((x_1+w)-x_1)^{n-2}dx_1$

=$n(n-1)w^{n-2}(1-w)$ (0 < w < 1)

wの式ですが、x_1で積分する点に注意しましょう。

また、$x_n = x_1 + w$ < 1より x_1 < 1-wで積分します。

② (2)の解法

期待値 E[R]は

E[R] = $\int_{0}^{1} r f_R(r)dr$

=$n(n-1)\int_{0}^{1} r^{n-1}(1-r)dr$

= $\frac{n-1}{n+1}$

(計算してみてください。)

となり、1に近い値になります。不思議！

③ (3)の解法

確率は確率密度関数$f_R(x)$を積分した結果になります。

Pr(R ≥ 1/2) = $\int_{1/2}^{1} f_R(r)dr$

=$n(n-1)\int_{1/2}^{1} r^{n-2}(1-r)dr$

=1-$(\frac{1}{2})^n(n+1)$

となります。

以上、「順序統計量の演習問題」を解説しました。