

## بازه‌های اطمینان بر اساس میانگین و دامنه

فاطمه یوسف‌زاده<sup>۲</sup>

ناصررضا ارقامی<sup>۱</sup>

### چکیده

گاهی اوقات پیش می‌آید که ما فقط بخشی از اطلاعات نمونه‌ای مشاهده شده را نگاه می‌داریم. در این گونه موارد تحلیل اطلاعات موجود از طریق روشهای آماری کلاسیک و متداول میسر نیست و بسته به اینکه چه اطلاعاتی از نمونه‌های مختلف در دست است، شیوه‌های خاص تحلیل مورد نیاز خواهد بود. ما در این مقاله این مسأله را مورد توجه قرار داده‌ایم که بر اساس مقادیر میانگین، ماکزیمم و مینیمم از تعداد  $m$  نمونه تصادفی از یک جامعه نرمال هر یک به حجم  $n$  چگونه می‌توان بازه اطمینان دقیق برای میانگین به دست آورد و این موضوع را در دو حالت  $m=1$  و  $m>1$  مورد بررسی قرار داده‌ایم. نتیجه کار تعدادی جدول است که به کمک آنها می‌توان به هدف فوق دست یافت. در این جداول ضریب اطمینان واقعی زمانی که توزیع جامعه نرمال نیست نیز ارائه شده است.

**واژه‌های کلیدی:** ماکسیمم نمونه، مینیمم نمونه، چندک، ضریب اطمینان، تقارن، شبیه‌سازی.

### ۱. حالت $m=1$

فرض کنید  $X_1, X_2, \dots, X_n$  یک نمونه تصادفی از توزیع  $N(\mu, \sigma^2)$  (هر دو مجهول) باشد و ما فقط به مقادیر  $\bar{x}, \max, \min$  از مشاهدات دسترسی داریم.

**قضیه:**  $t = \frac{\bar{x} - \mu}{r}$  که در آن  $r = \max - \min$  دامنه نمونه است، یک کمیت محوری متقارن است.

**اثبات:** فرض کنید  $z_i = \frac{x_i - \mu}{\sigma}$ ،  $i = 1, 2, \dots, n$ . در این صورت داریم؛

$$\bar{x} = \sigma \times \bar{z} + \mu \quad r = \sigma \times r_z$$

که  $\bar{z}$  و  $r_z$  به ترتیب میانگین و دامنه  $z_i$  ها هستند. می‌توان نوشت

$$t = \frac{(\sigma \times \bar{z} + \mu) - \mu}{\sigma \times r_z} = \frac{\bar{z}}{r_z}$$

که چون توزیع  $z_i$  ها به  $\mu$  و  $\sigma^2$  بستگی ندارند، توزیع  $t$  مستقل از پارامترهای مجهول می‌باشد.

برای اثبات تقارن توزیع  $t$  کافی است ثابت کنیم  $t$  و  $-t$  هم توزیع هستند. فرض کنید

$$(y_i - \mu) = -(x_i - \mu) \quad i = 1, 2, \dots, n$$

در این صورت داریم،

$$r_y = r \quad \bar{y} = -\bar{x} + 2\mu$$

در نتیجه

$$-t = \frac{-\bar{x} + \mu}{r} = \frac{(\bar{y} - 2\mu) + \mu}{r_y} = \frac{\bar{y} - \mu}{r_y} = t_y$$

<sup>۱</sup> گروه آمار، دانشگاه فردوسی مشهد

<sup>۲</sup> دانشجوی کارشناسی آمار، دانشگاه فردوسی مشهد

$$f(r, \bar{x}) = \frac{n(n-1)}{2\pi\sigma^2} \int_{-\infty}^{+\infty} [\Phi(\frac{x_{(1)} + r - \mu}{\sigma}) - \Phi(\frac{x_{(1)} - \mu}{\sigma})]^{n-2} \times e^{-\frac{1}{\sigma^2}[(x_{(1)} - \mu + \frac{r}{2})^2 + \frac{r^2}{4}]} dx_{(1)} \times \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \times e^{-\frac{n}{2\sigma^2}(\bar{x} - \mu)^2}$$

با انتخاب  $u = \bar{x} - \mu$ ،  $t = \frac{\bar{x} - \mu}{r}$  توزیع مشترک  $t, u$  را محاسبه می‌کنیم.

ژاکوبین این تبدیل به صورت زیر به دست می‌آید:

$$J(\bar{x}, r) = \frac{\bar{x} - \mu}{r^2}$$

$$f(t, u) = f(\bar{x}, r) \times \frac{r^2}{x - \mu} = f(u, \frac{u - \mu}{t}) \times \frac{u - \mu}{t^2}$$

حال توزیع کناری  $t$  عبارتست از:

$$f(t) = \int_{-\infty}^{+\infty} f(t, u) du = \frac{n\sqrt{n}(n-1)}{2\pi\sqrt{2\pi}\sigma^2} \times \int \int \frac{u - \mu}{t^2} e^{-\frac{1}{\sigma^2}[(x_{(1)} - \mu + \frac{u - \mu}{t})^2 + \frac{(u - \mu)^2}{t^2}]} \times [\Phi(\frac{x_{(1)} - \mu}{\sigma} + \frac{u - \mu}{t\sigma}) - \Phi(\frac{x_{(1)} - \mu}{\sigma})]^{n-2} \times e^{-\frac{n}{2\sigma^2}(u - \mu)^2} dx_{(1)} du$$

گرچه همانطور که ملاحظه می‌شود توزیع  $t$  به صورت بسته حاصل نمی‌شود، می‌توان (با توجه به اینکه توزیع  $t$  به  $\mu$  و  $\sigma^2$  بستگی ندارد چندکهای توزیع  $t$  را با استفاده از شبیه سازی به دست آورده و در جهت ساختن بازه اطمینان برای میانگین جامعه به کار برد. جدول (۱) چندکهای  $t$  برای حجم نمونه ۱۰ و ۲۰ را نشان می‌دهد. این مقادیر بوسیله شبیه سازی با تکرار ۱۰۰۰۰۰ بار انجام گرفته است.

اگر  $q_\alpha$  چندک توزیع  $t$  متناظر با  $1 - \alpha$  باشد با توجه به تقارن توزیع  $t$  می‌توان نوشت،

از طرفی چون (به خاطر تقارن توزیع نرمال حول میانگین جامعه)  $n_i$  ها و  $i$  ها هم توزیع اند،  $t_y$  و  $t$  هم توزیع می‌باشند. پس  $-t$  و  $t$  هم توزیع هستند.

متأسفانه توزیع این کمیت محوری به صورت بسته به دست نمی‌آید. با این حال می‌توان توزیع آن را به صورت زیر به دست آورد.

می‌دانیم که (زمانی که  $\sigma$  معلوم و  $\mu$  مجهول است)  $T = \bar{x}$  یک آماره بسنده کامل است و در نتیجه از  $r$  که یک آماره فرعی است (بنابر قضیه باسو) مستقل می‌باشد. برای اینکه بینیم  $r$  یک آماره فرعی است کافی است توجه کنیم که توزیع  $u_i$  ها ( $u_i = x_i - \mu$ )،  $i = 1, 2, \dots, n$ ، به  $\mu$  بستگی ندارد و  $r = r_u$ .

حال اگر  $F$  تابع توزیع  $r$  باشد، بنابراین (بنابر فصل ۶ نخستین درس احتمال. تألیف: اس. راس، ترجمه دکتر آذرنوش، دکتر بزرگ نیا، دکتر مشکانی و دکتر بیرومند).

$$F(r) = n \int_{-\infty}^{+\infty} [F(x_{(1)} + r) - F(x_{(1)})]^{n-1} f(x_{(1)}) dx_{(1)}$$

$$F(r) = n \int_{-\infty}^{+\infty} [\Phi(\frac{x_{(1)} + r - \mu}{\sigma}) - \Phi(\frac{x_{(1)} - \mu}{\sigma})]^{n-1} \times \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x_{(1)} - \mu)^2} dx_{(1)}$$

و از آنجا که

$$f(r) = \frac{\partial}{\partial r} F(r)$$

داریم

$$f(r) = \frac{n(n-1)}{2\pi\sigma^2} \int_{-\infty}^{+\infty} [\Phi(\frac{x_{(1)} + r - \mu}{\sigma}) - \Phi(\frac{x_{(1)} - \mu}{\sigma})]^{n-2} \times e^{-\frac{1}{\sigma^2}[(x_{(1)} - \mu + \frac{r}{2})^2 + \frac{r^2}{4}]} dx_{(1)}$$

دانستیم که  $r, \bar{x}$  مستقلند. بنابراین،

$$f(r, \bar{x}) = f(r) \times f(\bar{x})$$

**کمیت ۲:**

$$t = \frac{\bar{x} - \mu}{r}$$

به سادگی ثابت می‌شود که  $t$  یک کمیت محوری و متقارن است. چند کهای توزیع این کمیت در جدول ۵ ارائه شده است.

$$p(-q_{\frac{\alpha}{2}} < \frac{\bar{x} - \mu}{r} < q_{\frac{\alpha}{2}}) = 1 - \alpha$$

بازه اطمینان  $(1 - \alpha) \cdot 100$  درصد برای  $\mu$  عبارتست از،

$$(\bar{x} - q_{\frac{\alpha}{2}} \times r, \bar{x} + q_{\frac{\alpha}{2}} \times r)$$

در جدول ۶ ضرایب اطمینان واقعی و متوسط طول بازه اطمینان برای توزیع‌های مختلف به دست آمده است.

**کمیت ۳:**

$$t = \frac{\bar{x} - \mu}{R}, \quad R = Max - Min$$

در این کمیت  $Max$  ماکزیمم تمام نمونه‌ها و  $Min$  مینیمم تمام نمونه‌ها است. چند کهای این کمیت به صورت جدول ۷ است. بازه اطمینان به صورت زیر به دست می‌آید؛

$$p(-q_{\frac{\alpha}{2}} < \frac{\bar{x} - \mu}{R} < q_{\frac{\alpha}{2}}) = 1 - \alpha$$

بازه اطمینان  $(1 - \alpha) \cdot 100$  درصد برای  $\mu$  عبارتست از

$$s(\bar{x} - q_{\frac{\alpha}{2}} \times R, \bar{x} + q_{\frac{\alpha}{2}} \times R)$$

جدول ضرایب اطمینان واقعی و متوسط طول بازه اطمینان ۹۰٪ به صورت جدول ۸ خواهد بود.

**کمیت ۴:**

$$t = \frac{\bar{x} - \mu}{med}, \quad \text{میان } r_i \text{ ها } = med$$

چند کهای  $t$  با استفاده از توزیع نرمال استاندارد به صورت جدول ۹ به دست آمده است؛

$$p(-q_{\frac{\alpha}{2}} < \frac{\bar{x} - \mu}{Med} < q_{\frac{\alpha}{2}}) = 1 - \alpha$$

بازه اطمینان  $(1 - \alpha) \cdot 100$  درصد برای  $\mu$  عبارتست از،

$$p(-q_{\frac{\alpha}{2}} < \frac{\bar{x} - \mu}{r} < q_{\frac{\alpha}{2}}) = 1 - \alpha$$

در نتیجه یک بازه اطمینان  $(1 - \alpha) \cdot 100$  درصد برای  $\mu$  عبارتست از

$$(\bar{x} - q_{\frac{\alpha}{2}} \times r, \bar{x} + q_{\frac{\alpha}{2}} \times r)$$

جدول ۲ ضریب اطمینان واقعی و متوسط طول بازه اطمینان ۹۰٪ را برای توزیع‌های مختلف نشان می‌دهد.

**۲. حالت  $m > 1$**

فرض کنید  $(\bar{x}_1, \min_1, \max_1), \dots, (\bar{x}_m, \min_m, \max_m)$  را از مشاهدات  $m$  نمونه تصادفی مستقل  $\Pi$  تایی داریم. در این حالت چندین کمیت محوری می‌توان تصور کرد. هر کدام را به طور جداگانه مورد بررسی قرار می‌دهیم.

**کمیت ۱:**

$$t = \frac{1}{m} \sum_{i=1}^m \frac{\bar{x}_i - \mu}{r_i}$$

در اینجا نیز مانند قبل ثابت می‌شود که  $t$  یک کمیت محوری است و به پارامتر جامعه نیز بستگی ندارد. و نیز ثابت می‌شود که توزیع  $t$  متقارن است.

با استفاده از شبیه‌سازی توزیع نرمال استاندارد برای  $\Pi, m$  های مختلف چند کهای توزیع این کمیت را به دست آورده و در جدول ۳ قرار می‌دهیم.

$$t = \frac{1}{m} \sum_{i=1}^m \frac{\bar{x}_i - \mu}{r_i} = \frac{1}{m} \sum_{i=1}^m \frac{\bar{x}_i}{r_i} - \mu \times \frac{1}{m} \sum_{i=1}^m \frac{1}{r_i}$$

$$t = A - \mu \times B$$

$$p(-q_{\frac{\alpha}{2}} < A - \mu \times B < q_{\frac{\alpha}{2}}) = 1 - \alpha$$

بازه اطمینان  $(1 - \alpha) \cdot 100$  درصد برای  $\mu$  عبارتست از

$$\left( \frac{A - q_{\frac{\alpha}{2}}}{B}, \frac{A + q_{\frac{\alpha}{2}}}{B} \right)$$

جدول ۴ ضریب اطمینان واقعی و متوسط طول بازه اطمینان ۹۰٪ را نشان می‌دهد.

$$\bar{x} - \frac{(Med(max) - Med(min)) \times q_{\frac{\alpha}{2}}}{2}$$

$$\bar{x} - \frac{q_{\frac{\alpha}{2}} \times Med}{2}, \bar{x} + \frac{q_{\frac{\alpha}{2}} \times Med}{2}$$

$$\bar{x} + \frac{(Med(max) - Med(min)) \times q_{\frac{\alpha}{2}}}{2}$$

جدول ۱۰ ضرایب اطمینان واقعی و متوسط طول بازه اطمینان ۹۰٪ را نشان می‌دهد.

جدول ضرایب اطمینان واقعی و متوسط طول بازه اطمینان ۹۰٪ برای توزیع‌های مختلف در جدول ۱۲ آمده است.

کمیت ۵:

$$t = \frac{\bar{x} - \mu}{Med(max) - Med(min)}$$

چندکهای t در جدول ۱۱ آمده است؛

$$p\left(-q_{\frac{\alpha}{2}} < \frac{\bar{x} - \mu}{Med(max) - Med(min)} < q_{\frac{\alpha}{2}}\right) = 1 - \alpha$$

بازه اطمینان  $(1 - \alpha) \cdot 100$  درصد برای  $\mu$  عبارتست از،

### ۳. نتیجه‌گیری

با توجه به جداول بالا به نظر می‌رسد ضریب اطمینان واقعی برای توزیع‌های متقارن بیش از توزیع‌هایی است که دارای چولگی هستند. همچنین برای توزیع‌های متقارن که در آنها دم توزیع ضخیمتر از نرمال است ضریب اطمینان واقعی بیش از نرمال است. حتی برای توزیع‌های نامتقارن زمانی که  $m, n$  بزرگ باشد. ضریب اطمینان بیش از ضریب اطمینان اسمی خواهد بود.

جدول ۱

n	$1 - \alpha$	۰/۰۰۵	۰/۰۱	۰/۰۲۵	۰/۰۵	۰/۱
۱۰		۰/۳۳	۰/۲۹۵	۰/۲۳	۰/۱۸۵	۰/۱۳۸
۲۰		۰/۱۷۵	۰/۱۵۵	۰/۱۲۶	۰/۱۰۴	۰/۰۷۸

جدول ۲

n		$N(0,1)$	$E(1)$	$\chi^2_{(1)}$	$U(0,1)$	$Lo(0,1)$	$L(0,1)$	$C(0,1)$
۱۰	Real Confidence Coefficient	۰/۹۰	۰/۱۸۴	۰/۱۸۴	۰/۱۸۸	۰/۹۰	۰/۹۱	۰/۹۴
۱۰	Average Length	۱/۱۳۹	۱/۰۴۶	۱/۰۴۶	۰/۳۰۳	۲/۰۹۰	۱/۶۵۴	۳۵/۶۶۳
۲۰	Real Confidence Coefficient	۰/۹۰	۰/۱۸۶	۰/۱۸۶	۰/۱۸۵	۰/۹۱	۰/۹۲	۰/۹۶
۲۰	Average Length	۰/۷۷۷	۰/۷۳۷	۰/۷۳۷	۰/۱۸۸	۱/۴۷۷	۱/۲۰۶	۴۶/۱۱۹

جدول ۳

m	1-α		.005	.01	.025	.05	.1
	n						
5	10		.139	.130	.102	.1085	.1066
	20		.1076	.1067	.10565	.10477	.10367
10	10		.100	.1088	.10747	.10625	.1048
	20		.1067	.1061	.1052	.10419	.1032

جدول ۴

m	n		$N(0,1)$	$E(1)$	$\chi^2_{(1)}$	$U(0,1)$	$Lo(0,1)$	$L(0,1)$	$C(0,1)$
5	10	Real Confidence Coefficient	.190	.176	.174	.189	.192	.193	.194
		Average Length	.1494	.1416	.1417	.1137	.1887	.1679	1/978
5	20	Real Confidence Coefficient	.190	.183	.180	.188	.193	.193	.196
		Average Length	.1345	.1308	.1308	.1093	.1643	.1518	2/220
10	10	Real Confidence Coefficient	.190	.169	.165	.191	.193	.191	.195
		Average Length	.1361	.1300	.1301	.1100	.1645	.1489	1/330
10	20	Real Confidence Coefficient	.190	.188	.182	.197	.198	.199	.199
		Average Length	.1301	.1268	.1268	.1047	.1556	.1446	1/774

جدول ۷

m	1-α		.005	.01	.025	.05	.1
	n						
۵	۱۰		-.109	-.108	-.1065	-.1054	-.1042
	۲۰		-.1057	-.10495	-.1041	-.1034	-.10257
۱۰	۱۰		-.1056	-.1051	-.1042	-.1035	-.1027
	۲۰		-.1049	-.1042	-.1036	-.1029	-.1022

جدول ۸

m	n		$N(0,1)$	$E(1)$	$\chi^2_{(1)}$	$U(0,1)$	$Lo(0,1)$	$L(0,1)$	$C(0,1)$
۵	۱۰	Real Confidence Coefficient	.190	.189	.187	.180	.193	.195	.199
۵	۱۰	Average Length	.1485	.1484	.1484	.1104	.1969	.1824	۲۶/۳۰۹
۵	۲۰	Real Confidence Coefficient	.190	.191	.186	.174	.193	.195	.199
۵	۲۰	Average Length	.1340	.1353	.1353	.1067	.1704	.1611	۴۷/۰۰۲
۱۰	۱۰	Real Confidence Coefficient	.190	.192	.193	.179	.195	.197	.1998
۱۰	۱۰	Average Length	.1351	.1363	.1363	.1064	.1725	.1631	۵۵/۶۳۶
۱۰	۲۰	Real Confidence Coefficient	.190	.198	.195	.190	.199	.1999	۱
۱۰	۲۰	Average Length	.1303	.1340	.1340	.1043	.1682	.1591	۸۷/۲۱۹

جدول ۹

m	$1-\alpha$		.005	.01	.025	.05	.1
	n						
۵	۱۰		-.114	-.1118	-.10955	-.108	-.1062
	۲۰		-.1076	-.1067	-.1055	-.1046	-.1035
۱۰	۱۰		-.109	-.1083	-.107	-.10575	-.1044
	۲۰		-.1069	-.1059	-.1049	-.1040	-.103

جدول ۱۰

m	n		$N(0,1)$	$E(1)$	$\chi^2_{(1)}$	$U(0,1)$	$Lo(0,1)$	$L(0,1)$	$C(0,1)$
۵	۱۰	Real Confidence Coefficient	.190	.186	.186	.190	.192	.192	.152
۵	۱۰	Average Length	.1485	.1425	.1425	.1333	.1878	.1687	2/598
۵	۲۰	Real Confidence Coefficient	.190	.187	.180	.185	.192	.192	.157
۵	۲۰	Average Length	.1343	.1314	.1314	.1096	.1644	.1522	2/829
۱۰	۱۰	Real Confidence Coefficient	.190	.188	.185	.192	.192	.192	.141
۱۰	۱۰	Average Length	.1350	.1303	.1305	.1096	.1630	.1489	1/660
۱۰	۲۰	Real Confidence Coefficient	.190	.192	.191	.197	.198	.197	.152
۱۰	۲۰	Average Length	.1300	.1269	.1269	.1073	.1552	.1448	2/325

جدول ۱۱

m	1-α		.005	.01	.025	.05	.1
	n						
5	10		-.113	-.112	-.1099	-.1080	-.10623
	20		-.1078	-.107	-.1056	-.1046	-.10354
10	10		-.109	-.1083	-.1068	-.1057	-.1045
	20		-.1068	-.1059	-.1051	-.1041	-.10307

جدول ۱۲

m	n		$N(0,1)$	$E(1)$	$\chi^2_{(1)}$	$U(0,1)$	$Lo(0,1)$	$L(0,1)$	$C(0,1)$
5	10	Real Confidence Coefficient	.190	.186	.186	.192	.191	.190	.145
5	10	Average Length	.1483	.1429	.1430	.1136	.1861	.1662	1/970
5	20	Real Confidence Coefficient	.190	.187	.184	.189	.192	.194	.150
5	20	Average Length	.1338	.1312	.1312	.1085	.1625	.1505	2/281
10	10	Real Confidence Coefficient	.190	.189	.182	.192	.191	.189	.133
10	10	Average Length	.1344	.1303	.1303	.1095	.1618	.1484	1/222
10	20	Real Confidence Coefficient	.190	.195	.193	.197	.198	.197	.143
10	20	Average Length	.1300	.1277	.1277	.1042	.1540	.1434	1/721

## مراجع

- [۱] پاریسیان، احمد، مبانی آمار ریاضی، دانشگاه صنعتی اصفهان، ۱۳۷۸.
- [۲] راس، شلدون ام، شبیه‌سازی، ترجمه دکتر حسنعلی آذرنوش، دکتر حسینعلی نیرومند، دانشگاه فردوسی مشهد، ۱۳۷۸.
- [۳] راس، شلدون ام، نخستین درس در احتمال، ترجمه دکتر حسنعلی آذرنوش، دکتر ابوالقاسم بزرگنیا، دانشگاه فردوسی مشهد، ۱۳۷۶.
- [۴] مود، الکساندر ام، گریبیل، فرانکلین آ. و بوز، دون اس، مقدمه‌ای بر نظریه آمار، ترجمه دکتر علی مشکاتی، دانشگاه فردوسی مشهد، ۱۳۷۷.