

استنباط آماری براساس روش برآوردگر گشتاوری وزنی احتمالی تعدیل یافته

افشین فیاض موقر^{*}، زهره محمودی^۲

۱- استادیار دانشگاه مازندران، دانشکده علوم ریاضی، گروه آمار، بابلسر، ایران

۲- دانشجوی کارشناسی ارشد، دانشگاه مازندران، گروه آمار، بابلسر، ایران

رسید مقاله: ۱۳۹۳ اسفند

پذیرش مقاله: ۱۳۹۴ مرداد

چکیده

در این مقاله روش جدیدی در نظریه برآوردهای با عنوان برآوردگر گشتاوری وزنی احتمالی تعدیل یافته (APWM)^۱ معرفی می‌گردد، که تعیینی از روش گشتاوری وزنی احتمالی (PWM)^۲ می‌باشد. با استفاده از این روش استنباط آماری بر روی پارامتر توزیع یکنواخت $(0, \theta)$ انجام داده‌ایم، و نتیجه آن را با نتایج روش‌های دیگر برآوردهای از جمله روش حداقل نتایج (ML)^۳ که برای توزیع یکنواخت دارای کوتاه‌ترین فاصله اطمینان می‌باشد و روش گشتاوری وزنی احتمالی (PWM) مقایسه نموده‌ایم. نتایج بیانگر برتری APWM نسبت به PWM می‌باشد و در اکثر مواقع حداقل نتایج مشابه ML می‌دهد.

کلمات کلیدی: برآوردگر گشتاوری وزنی احتمالی، مجموع اشتیل یس تصادفی.

۱ مقدمه

برآورد پارامترهای توزیع‌های احتمالی و معرفی برآوردگرها مناسب یکی از موضوع‌های مهم در آمار می‌باشد، که از جمله روش‌های برآوردهای می‌توان به روش گشتاوری (MM)^۴، روش حداقل نتایج (ML)^۵ و... اشاره کرد [۱].

براساس این برآوردگرها استنباط‌هایی در رابطه با پارامتر توزیع صورت می‌پذیرد. از جمله این استنباط‌ها می‌توان فاصله اطمینان و انجام آزمون را نام برد. مفاهیم برآورد فاصله‌ای و آزمون فرض‌ها ارتباط نزدیکی با هم دارند و هر یک از این مفاهیم می‌تواند منجر به معرفی مفهوم دیگر شود. همان‌طور که در برآورد نقطه‌ای معیار بررسی عملکرد روش‌های مختلف برآوردهایی، میانگین مربع خطأ^۶ می‌باشد در برآورد فاصله‌ای برای بررسی

* عهدہ دار مکاتبات

آدرس الکترونیکی: a_fayyaz@umz.ac.ir

² Adjusted Probability Weighted Moments

³ Probability Weighted Moments

⁴ Maximum Likelihood

⁵ Method of Moments

⁶ Mean Square Error

عملکرد، توان^۱ مربوط به روش‌ها را معیار قرار می‌دهیم. در این مقاله به بررسی عملکرد روش‌های مختلف برآوردهایی در انجام استنباط آماری می‌پردازیم.

در بخش دوم به معرفی، روش گشتاوری وزنی احتمال (PWM) و در بخش سوم روش جدید، گشتاوری وزنی احتمالی تبدیل یافته (APWM)، خواهیم پرداخت. در بخش چهارم استنباط‌های آماری مربوط به پارامتر توزیع یکنواخت ارایه می‌گردد و مقایسه عددی مربوط به استنباط‌های مختلف در بخش پنجم آمده است.

۲ روش برآوردهایی PWM

شاید بتوان گفت که روش گشتاوری یکی از قدیمی‌ترین روش‌ها برای برآورد پارامترها می‌باشد، که توسط کارل پیرسن در سال ۱۸۹۴ معرفی گردید. در سال ۱۹۷۹ گرین وود^[۲] با وزنی کردن روش گشتاوری (MM)، روش جدیدی را با عنوان روش گشتاوری وزنی احتمالی (PWM) معرفی کرد، که در واقع تعمیمی از روش گشتاوری می‌باشد.

گشتاور وزنی احتمالی یک متغیر تصادفی X ، باتابع توزیع $F_\theta(x) = P_\theta(X \leq x)$ به صورت زیر معرفی می‌گردد:

$$M_{(p,r,s)} = E_\theta \left[X^p \{F_\theta(X)\}^r \{1 - F_\theta(X)\}^s \right] \quad (1)$$

که در آن P , r و s مقادیر حقیقی می‌باشند. هنگامی که معکوستابع توزیع، $(F(x))$ را بتوان به فرم بسته‌ای نوشت، رابطه (1) به صورت زیر نوشته می‌شود:

$$M_{(p,r,s)} = \int \{x(F)\}^p F^r (1-F)^s dF \quad (2)$$

که در آن F دارای توزیع یکنواخت (۱) می‌باشد، که به این ترتیب گشتاورها وزنی احتمالی را می‌توان آسان‌تر به دست آورد. همان‌طور که مشاهده می‌شود کمیت‌های $M_{(p,.,s)}$ ($p = 1, 2, \dots$) گشتاورهای معمولی غیر مرکزی به دست می‌دهند. آماره برآوردگر پارامتر θ مشابه روش برآوردهایی گشتاوری، از برابری گشتاورهای وزنی احتمالی جامعه (۱) با مقدار به دست آمده آن براساس نمونه حاصل می‌گردد. در حالت خاص که $s = 0$ یا $r = 0$ فرض می‌گردد [۳] و برآوردگر پارامتر θ از رابطه‌های زیر حاصل می‌شود:

$$a_s = M_{(1,.,s)} = E_\theta \left[X \{1 - F_\theta(X)\}^s \right] \quad (3)$$

یا

$$b_r = M_{(1,r,.)} = E_\theta \left[X \{F_\theta(X)\}^r \right] \quad (4)$$

که در آن

^۱ Power

$$a_s = \frac{1}{n} \sum_{j=1}^n \frac{\binom{n-j}{r}}{\binom{n-1}{r}} x_{(j)}, \quad (5)$$

و

$$b_r = \frac{1}{n} \sum_{j=1}^n \frac{\binom{j-1}{r}}{\binom{n-1}{r}} x_{(j)} \quad (6)$$

و $x_{(j)}$ مقدار مشاهده شده (j) ، X ، j -امین آماره ترتیبی نمونه تصادفی n تایی می‌باشد. از این روش در برآورد پارامترهای چندین توزیع مهم از جمله گامبل، لجستیک و وایل استفاده شده است که نتایج مطلوبی نسبت به سایر روش‌های برآوردهای ارایه نموده است [۳] و [۴]. همچنین از خواص این برآوردگر می‌توان به مجانبی نرم‌افزار آن اشاره کرد [۴].

۳ معرفی روش جدید برآوردهای APWM

در این بخش، روشی جدید در برآوردهای پارامتر جامعه با عنوان، برآورد گشتاوری وزنی احتمالی تعدیل یافته (APWM) معرفی می‌گردد که در واقع تعمیم یافته روش گشتاوری وزنی احتمالی می‌باشد. این ایده از روش گشتاوری تعدیل یافته (AMM)^۱، که توسط سلطانی و حومه‌ای [۵] در سال ۲۰۰۹ معرفی گردیده، گرفته شده است. روش جدید مشابه روش PWM می‌باشد، اما به جای استفاده از گشتاورهای وزنی احتمالی نمونه از مجموع جزئی استلت‌جس تصادفی (RSPS)^۲ متناظر استفاده می‌گردد.

فرض کنید X_1, X_2, \dots, X_n یک نمونه تصادفی n تایی از جامعه‌ای با تابع توزیع دلخواه F_θ با پارامتر مجهول θ باشد. همان‌طور که در بخش ۱ به آن اشاره شد گشتاور وزنی احتمالی را می‌توان به صورت زیر نوشت:

$$M_{(p,r,s)} = E_\theta \left[X^p \{F_\theta(X)\}^r \{1 - F_\theta(X)\}^s \right] = \int \{x(F)\}^p F^r (1-F)^s dF \quad (7)$$

بر اساس نظریه انتگرال گیری کلاسیک، انتگرال موجود در رابطه بالا را می‌توان از مجموع استلت‌جس تصادفی بالایی (پایینی) برای $p, r, s = 0, 1, 2, \dots$ به صورت زیر تقریب زد:

$$\begin{aligned} X_{(U)(p,r,s)} &= \sum_{j=1}^n X_{(j)}^p \{F(X_{(j)})\}^r \{1 - F(X_{(j)})\}^s \left[F(X_{(j)}) - F(X_{(j-1)}) \right], \\ (X_{(L)F})^n &= \sum_{j=1}^n X_{(j)}^p \{F(X_{(j)})\}^r \{1 - F(X_{(j)})\}^s \left[F(X_{(j+1)}) - F(X_{(j)}) \right] \end{aligned} \quad (8)$$

که در آن $F(X_{(n+1)}) = 1$ ، $F(X_{(1)}) = 0$ و $X_{(1)} \leq \dots \leq X_{(n)}$ آماره متناظر با X_1, X_2, \dots, X_n می‌باشد. در این مقاله ما برآورد پارامتر مجهول θ را از حل دستگاه غیر خطی

¹ Adjusted Method of Moments
² Random Stieltjes Partial Sums

$$M_{(p,r,s)} = X_{(U)(p,r,s)} \quad (9)$$

برآورده می‌کنیم.

فیاض و محمودی [۶] را برای توزیع یکنواخت $(0,\theta)$ به دست آورده‌اند و مقایسه‌ای بین نتایج این روش و روش‌های حداکثر درستنمایی، گشتاوری تعديل یافته و گشتاوری وزنی احتمالی برای توزیع یکنواخت از طریق مقایسه میانگین مربع خطأ انجام داده اند که نشان داد روش APWM برتری‌هایی نسبت به بقیه روش‌ها دارد. برآورده APWM پارامتر توزیع یکنواخت $(0,\theta)$ به صورت زیر به دست می‌آید [۶]:

$$\hat{\theta}_{APWM} = \left[(r+2) \sum_{j=1}^n U_{(j)}^{(r+1)} (U_{(j)} - U_{(j-1)}) \right]^{\frac{1}{(r+2)}}$$

سلطانی و عبدالنژاد دو روش استنباط آماری را بر طبق برآورده‌گرهای گشتاوری تعديل یافته معرفی کرده‌اند [۷] و [۸].

در این مقاله هدف ما انجام استنباط آماری بر روی پارامتر مجھول توزیع یکنواخت می‌باشد، که در ادامه به این موضوع خواهیم پرداخت.

۴ استنباط برای پارامتر توزیع یکنواخت

سلطانی و عبدالنژاد با استفاده از برآورده‌گرهای روش AMM بر روی پارامتر θ توزیع یکنواخت، استنباط آماری انجام داده‌اند [۷] که ما در این بخش با استفاده از برآورده‌گرهای APWM استنباط آماری بر روی پارامتر توزیع یکنواخت انجام خواهیم داد.

در این بخش جهت استنباط در مورد پارامتر θ ، فاصله اطمینانی را برای آن با استفاده از APWM برآورده ساخته و آزمون فرضی بر روی آن انجام می‌گردد.

۴-۱ فاصله اطمینان

فرض کنید، U_1, \dots, U_n یک نمونه تصادفی از $(0,\theta)$ با پارامتر θ باشد و $U_{(1)}, \dots, U_{(n)}$ آماره‌های مرتب متناظر با این نمونه تصادفی باشند.

کوتاه‌ترین فاصله اطمینان $1 - \alpha$ ٪ حاصل از برآورده گر ML

$$\left(U_{(n)}, \frac{U_{(n)}}{\sqrt[n]{\alpha}} \right) \quad (10)$$

می‌باشد [۱].

به منظور به دست آوردن فاصله اطمینان APWM عبارت زیر را در نظر بگیرید:

$$B = \frac{\hat{\theta}_{APWM}}{\theta} = \left[(r+2) \sum_{j=1}^n \frac{U_{(j)}}{\theta} \left\{ \frac{U_{(j)}}{\theta} \right\}^r \left(\left(\frac{U_{(j)}}{\theta} \right) - \left(\frac{U_{(j-1)}}{\theta} \right) \right) \right]^{\frac{1}{(r+2)}}$$

¹ Shortest Confidence Interval

از آنجا که توزیع عبارت بالا مستقل از θ می‌باشد بنابراین به عنوان کمیت محوری^۱ در نظر گرفته می‌شود.
بنابراین فاصله اطمینان $100(1-\alpha)\%$ پارامتر θ براساس کمیت محوری B به صورت زیر محاسبه می‌گردد.
می‌دانیم که

$$P\left(B_{\left(\frac{\alpha}{r}\right)} \leq \frac{\hat{\theta}_{APWM}}{\theta} \leq B_{\left(1-\frac{\alpha}{r}\right)}\right) = 1-\alpha,$$

که در آن $B_{\left(1-\frac{\alpha}{r}\right)}$ و $B_{\left(\frac{\alpha}{r}\right)}$ چندک‌های متغیر تصادفی B می‌باشند.

در نتیجه داریم:

$$\begin{aligned} 1-\alpha &= P\left(B_{\left(\frac{\alpha}{r}\right)} \leq \frac{\hat{\theta}_{APWM}}{\theta} \leq B_{\left(1-\frac{\alpha}{r}\right)}\right) = P\left(\frac{\left[(r+1)\sum_{j=1}^n u_{(j)}^{(r+1)}(u_{(j)} - u_{(j-1)})\right]^{1/(r+1)}}{B_{\left(1-\frac{\alpha}{r}\right)}}\right. \\ &\leq \theta \leq \left.\frac{\left[(r+1)\sum_{j=1}^n u_{(j)}^{(r+1)}(u_{(j)} - u_{(j-1)})\right]^{1/(r+1)}}{B_{\left(\frac{\alpha}{r}\right)}}\right). \end{aligned}$$

متغیر تصادفی T را به صورت زیر تعریف می‌کنیم:

$$T = \frac{\left[(r+1)\sum_{j=1}^n u_{(j)}^{(r+1)}(u_{(j)} - u_{(j-1)})\right]^{1/(r+1)}}{B} = \left(\frac{\sum_{j=1}^n u_{(j)}^{(r+1)}(u_{(j)} - u_{(j-1)})}{\sum_{j=1}^n V_{(j)}^{(r+1)}(V_{(j)} - V_{(j-1)})}\right)^{1/(r+1)} \quad (11)$$

که در آن $u_{(j)}$ مقدار مشاهده شده، $V_{(j)}$ می‌باشد (علامت d نشان دهنده هم توزیع بودن^۲ است).

حال فاصله اطمینان $100(1-\alpha)\%$ برای پارامتر θ را می‌توان به صورت

$$\left(T_{\left(\frac{\alpha}{r}\right)}, T_{\left(1-\frac{\alpha}{r}\right)}\right)$$

به دست آورد که در آن $T_{(\alpha)}$ چندک ام توزیع T می‌باشد. در عمل چندک‌های T از طریق شبیه سازی به دست می‌آیند.

۴-۲ آزمون فرض

برای انجام آزمون فرض برای پارامتر θ ، سه نوع آزمون می‌توانیم انجام دهیم:

(i) $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$,

(ii) $H_0 : \theta \leq \theta_0$ vs $H_1 : \theta > \theta_0$,

¹ Pivotal Quantity

² Identically Distributed

(iii) $H_0: \theta \geq \theta_0$ vs $H_1: \theta < \theta_0$.

در روش ML ناحیه بحرانی مربوط به پرتوان ترین آزمون یکنواخت، فواصل راست و چپ فاصله اطمینان بیان شده در رابطه (۱۰) می‌باشد، یعنی

$$(-\infty, U_{(n)}) \cup \left(\frac{U_{(n)}}{\sqrt[n]{a}}, \infty \right),$$

- در روش APWM نیز ناحیه رد هر یک از آزمون‌های بالا بر حسب متغیر تصادفی T به ترتیب عبارتند از:
- i. $|T| \geq \theta_0$.
 - ii. $T \leq \theta_0$.
 - iii. $T \geq \theta_0$.

بنابراین P -مقدار^۱ این آزمون‌ها که آن‌ها را با P نمایش می‌دهیم، برابرند با:

- i. $P = \min\{P(T \leq \theta_0), P(T \geq \theta_0)\}$
- ii. $P = P(T \leq \theta_0)$
- iii. $P = P(T \geq \theta_0)$

۵ مثال عددی

در این بخش محاسبات عددی مربوط به نتایج حاصل شده در بخش ۴ براساس شبیه سازی به دست می‌آید. در ادامه این نتایج با روش‌های PWM و ML به صورت عددی مقایسه می‌گردد.

۱-۵ بررسی فاصله اطمینان

جهت مقایسه میانگین طول^۲ و احتمال پوشش^۳ فواصل اطمینان حاصل از برآوردگر APWM و PWM (کوتاه‌ترین فاصله اطمینان) از شبیه سازی استفاده می‌شود. الگوریتم زیر مراحل شبیه سازی متغیر تصادفی T را در چهار گام نشان می‌دهد:

۱. تولید U_i ها از $U(\cdot, \theta_0)$ ، $i = 1, 2, \dots, n$.
۲. تولید V_i ها از $V(\cdot, 1)$ ، $i = 1, 2, \dots, n$.
۳. محاسبه متغیر T با استفاده از رابطه (۱۱).
۴. m بار تکرار سه مرحله قبل.

جدول (۱) نتایج مربوط به این مقایسه به ازای n های متفاوت، $r = 2$ ، $m = 1000$ و $\alpha = 0.05$ می‌باشد، که بیانگر آن است که احتمال پوشش برای هر اندازه نمونه، نزدیک به سطح ۹۵٪ می‌باشد و روش APWM به خوبی عمل کرده است.

¹ P-value

² Length Average

³ Coverage Probability

کمترین احتمال پوشش در بین این سه روش متعلق به روش PWM می‌باشد. همچنین برای هر n ، APWM دارای میانگین طول کمتری نسبت به PWM می‌باشد.

جدول ۱. احتمال پوشش (C) و میانگین طول (L) فاصله اطمینان دو طرفه ۹۵٪ برای پارامتر توزیع یکنواخت با $\theta = 2$ با توجه به حجم های نمونه مختلف

روش \ نمونه (n)	۵	۱۰	۳۰	۵۰	۷۰	۱۰۰	۵۰۰
APWM							
C	۰/۹۵	۰/۹۵	۰/۹۶	۰/۹۵	۰/۹۶	۰/۹۵	۰/۹۵
L	۱/۲۱۸۶	۰/۷۵۳۰	۰/۲۲۸۲	۰/۱۷۳۷	۰/۱۱۰۳	۰/۰۷۵۵	۰/۰۴۳۲
(PWM)							
C	۰/۸۷	۰/۸۸	۰/۹۲	۰/۹۵	۰/۹۴	۰/۹۵	۰/۹۵
L	۲/۲۵۸۵	۱/۲۳۵۷	۰/۶۱۶۶	۰/۴۳۲۲	۰/۳۷۲۷	۰/۲۶۸۴	۰/۱۱۳۳
(ML)							
C	۰/۹۵	۰/۹۶	۰/۹۵	۰/۹۶	۰/۹۶	۰/۹۵	۰/۹۵
L	۱/۴۷۲۳	۰/۶۷۶۴	۰/۲۰۴۸	۰/۱۲۲۲	۰/۰۸۵۶	۰/۰۵۹۹	۰/۰۳۲۱

۲-۵ نتایج مربوط به آزمون فرض

یک مطالعه شبیه‌سازی جهت مقایسه اندازه^۱ و توان آزمون براساس برآوردگر APWM، PWM و ML صورت گرفته است. بدین منظور نمونه‌هایی با اندازه‌های نمونه متفاوت (n) از توزیع (θ, r) تولید کرده و با ۵۰۰۰ بار تکرار این عمل، توان و اندازه آزمون را به دست می‌آوریم. نتایج عددی برای $\theta = 2$ در جداول ۲ و ۳ گردآوری شده‌اند.

در بیشتر مواقع، اندازه APWM در بین سه روش دیگر بهتر و همیشه نزدیک به سطح $\alpha = 0/05$ می‌باشد (جدول ۲). جدول (۳) توان آزمون را برای فرض (*i*) نشان می‌دهد. همان طور که مشاهده می‌گردد در سطح APWM دارای توان بزرگتری می‌باشد و PWM نتایج ضعیف‌تری را نشان می‌دهد، در حالی که $\alpha = 0/05$ APWM و ML به یکدیگر نزدیک می‌باشند.

جدول ۲. اندازه آزمون در آزمون دو طرفه $\alpha = 0/05$ توزیع یکنواخت به ازای $\theta = 2$ برای حجم های نمونه متفاوت

روش \ n	۵	۱۰	۳۰	۵۰	۷۰	۱۰۰	۵۰۰
APWM	۰/۰۵	۰/۰۵	۰/۰۴	۰/۰۵	۰/۰۴	۰/۰۵	۰/۰۵
PWM	۰/۱۳	۰/۱۲	۰/۰۸	۰/۰۵	۰/۰۶	۰/۰۵	۰/۰۵
ML	۰/۰۵	۰/۰۴	۰/۰۵	۰/۰۴	۰/۰۴	۰/۰۵	۰/۰۵

^۱ Size

جدول ۳. توان آزمون در آزمون دو طرفه $\alpha = 0.05$ توزیع یکنواخت به ازای $\theta_0 = 2$ برای θ ها و حجم‌های نمونه متفاوت

روش	نمونه (n)						
	5	10	30	50	70	100	500
$(\theta_0 = 1/5)$							
APWM	0.80	0.96	1	1	1	1	1
PWM	0.77	0.88	1	1	1	1	1
ML	0.73	0.95	1	1	1	1	1
$(\theta_0 = 1/75)$							
APWM	0.58	0.77	0.98	1	1	1	1
PWM	0.49	0.57	0.74	0.83	0.95	0.99	1
ML	0.62	0.84	0.98	1	1	1	1
$(\theta_0 = 2)$							
APWM	0.17	0.37	0.73	0.94	0.96	1	1
PWM	0.31	0.34	0.35	0.41	0.48	0.54	0.89
ML	0.27	0.50	0.78	0.96	0.95	1	1

۶ نتیجه‌گیری

در این مقاله ما بر پایه روش APWM یک فاصله اطمینان معرفی کردیم و از طریق شبیه سازی احتمال پوشش و میانگین طول روش APWM را با روش‌های ML و PWM مورد بررسی قرار دادیم.

برای حجم‌های متفاوت نمونه کوتاه‌ترین فاصله اطمینان مربوط به روش APWM بوده است. علاوه بر این احتمال پوشش به سطح اطمینان ۹۵٪ نزدیک می‌باشد. همچنین با توجه به نتایج عددی بخش آزمون فرض‌ها عملکرد بهتر روش APWM را نسبت به دو روش تحت بررسی نشان می‌دهد.

بنابراین با توجه به نتایج عددی مشاهده شده، می‌توان گفت که روش APWM به همان خوبی روش ML عمل می‌کند. این در حالی است که روش PWM ضعیف‌ترین نتایج را به دست می‌دهد.

منابع

[۱] فیاض موقر، ا، محمودی، ز، (۱۳۹۳). روش نوین برآوردگرهای گشتاوری با وزن احتمالی تعدیل یافته، دوازدهمین کنفرانس آمار ایران، کرمانشاه.

- [1] Casella, G., Berger, R. L., (1989). Statistics Inference, Wadsworth, Pacific Grove, CA.
- [2] Greenwood, J. A., Landwehr, J. M., Matalas, N. C., Wallis, J. R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. Water Resources Research, 15, 1049-1054.
- [3] Landwehr, J. M., Matalas, N. C., Wallis, J. R., (1979). Probability weighted moments compared with some traditional in estimation gumbel parameters and quantiles, Water Resources Research, 15, 1055–1064
- [4] Hosking, J. R. M., Wallis, J. R., Wood, E. F., (1985). Estimation of the generalized extreme value distribution by the method of probability weighted moments. Technometrics, 27, 251-261.
- [5] Soltani, A. R., Homei, H., (2009). A generalization for two-sided power distribution and adjusted method of moments. Statistics, 43, 611-620.

- [7] Soltani, A. R., Abdollahnezhad, K., (2011). On adjusted method of moments estimators on uniform distribution samples, Metron, 70, 27-40.
- [8] Soltani, A. R., Abdollahnezhad, K., (2012). A new computational test procedure adjusted method of moments, Journal of Applied Probability and Statistics, 7, 9-18.