بسم الله الرحمن الرحيم

دانشکده علوم

گروه آمار

پایان‌نامه جهت اخذ درجه‌ی کارشناسی ارشد رشته‌ی آمار گرایش محض

**عنوان پایان نامه**

# برآورد مینیماکس پارامتر محدود شده‌ی توزیع‌های گسسته تحت تابع زیان مربع خطا ولاینکس

استاد راهنما:

دکتر داوود قزوینی نژاد

نگارش:

محمد یاراحمدی

مهرماه ۱۳۸۸

**تقدیر و تشکر:**

از تمامی اساتید محترم گروه آمار ، بویژه استاد ارجمندم جناب آقای دکتر داوود قزوینی‌نژاد
که درکنار آموختن علم، درس گذشت، ایثار و خوب بودن را به این حقیر هدیه کرد.
در پناهش پیروز و سرفراز، در مهرش غوطه‌ور، به درگاهش نیازمند
و در لوایش همیشگی و ماندگار بمانید.

**تقدیم به:**

پدر و مادر و همسر مهربانم که با مهر و محبت مرا به سوی خدا رهنمون کرده و سعی و تلاش کردن، صبر کردن، رسیدن و سبزی زندگی به من آموختند.

به یاریش برایم همیشگی خواهید ماند.

## چکیده

در بسیاری از مسائل علمی و عملی؛ بویژه در مسائل پزشکی؛ گاهی اوقات صفت مورد بررسی عملاً در جامعه یا نادر و کمیاب و یا به‌گونه‌ای است که می‌دانیم به سمت معلومی میل می‌کند. هدف برآورد پارامتر مجهول تحت بررسی، بوسیله‌ی یکی از روشهای برآوردیابی است. در چنین مسائلی با اعمال محدودیت‌هایی بر روی فضای دامنه‌ی پارامتر مجهول؛ مسأله را حل می‌کنیم.
در آمار گاهی اوقات با چنین مسائلی مواجه می‌شویم، یعنی می‌بایست با در نظر گرفتن برخی قیود و محدودیت‌ها و اعمال آنها؛ پارامتر تحت بررسی را برآورد می‌نماییم. در این پایان‌نامه علاقه‌مند به دست‌یابی برآوردگر بیز و مینیماکس پارامتر مجهول برخی از توزیع‌های گسسته؛ تحت تابع زیان مربع خطا و لاینکس با در نظر گرفتن محدودیتی بر روی فضای پارامتر هستیم. با توجه به شرایط و قیود ذکر شده، می‌بایست یک توزیع پیشین دونقطه‌ای برای پارامتر تحت بررسی در نظر بگیریم، تا یک تصمیم بیز برابرکننده بدست آوریم. چون می‌خواهیم برآوردگر بیزی که بدست می‌آوریم؛ مینیماکس نیز باشد؛ می‌بایست این برآوردگر در چند شرط صدق کند. با بررسی این شرایط برآوردگرهای بیز و مینیماکس بدست آمده فرم ساده‌ای خواهند داشت. البته با لحاظ کردن این شرایط؛ بر روی فضای پارامتر تحت بررسی محدودیتی ایجاد می‌شود، که به ازای برخی از مقادیر فضای پارامتر، برآوردگرهای بیز بدست آمده، مینیماکس نیز هستند.

**فهرست مطالب**

**عنوان** **صفحه**

# فصل اول

## نظریه‌ی تصمیم

## ۱-۱ مقدمه

انسانهای مسئول و منطقی؛ آنها که دارای عقل سلیم هستند، با علم وآگاهی کافی؛ با درنظر گرفتن احتمال پیشامدها با توجه به سود و زیان و یا شهرت و رسوائی تصمیم می‌گیرند وکاری را انجام می‌دهند. دراین فصل برخی از این واژه‌ها را به عنوان اصطلاحات نظریه‌ی تصمیم، با دقت لازم شرح می‌دهیم.

در سالهای‌آخر نیمه‌ی اول قرن بیستم ابراهام والد، اصول تصمیم‌گیری را در چارچوبی ریاضیاتی ارائه داد و کوشید که مباحث گوناگون آماری مانند برآوردیابی، آزمونها و فاصله‌ی اطمینان را بهم پیوند دهد و از این راه نظریه‌ی آمار را پربارتر سازد. تمامی روش‌های استنباط آماری که تاکنون بررسی کرده‌ایم یک نوع تصمیم‌گیری می‌باشند، که در آنها به کمک نمونه‌های آماری و داده‌ها درباره‌ی پارامتر یا توزیع مجهول به تصمیم‌گیری می‌پرداختیم، این روش‌ها به آمار کلاسیک موسوم بودند. در نظریه‌ی تصمیم به کمک نمونه‌های آماری و دیگر جوانب همانند سود و زیان به دنبال دست‌یابی به بهترین تصمیم هستیم. بنابراین نظریه‌ی تصمیم شامل روش‌های دیگری از مسائل استنباط آماری است که روش‌های قبلی را شامل می‌شود و جنبه‌های جدیدی را نیز در نظر می‌گیرد. قبل از اینکه تابع تصمیم را تعریف کنیم چند اصطلاح مهم را شرح می‌دهیم.

### تعریف۱-۱-۱:

مقدار مجهولی که ممکن است در تصمیم‌گیری مؤثر باشد با $\theta$نمایش داده و آنرا وضع طبیعت می‌نامیم. از نماد $\Theta$ به عنوان مجموعه‌ی تمام $\theta$های ممکن استفاده می‌کنیم، بنابراین : $\theta \in \Theta$

مجموعه‌ی اوضاع طبیعت ممکن است متناهی و یا نامتناهی باشد، در آمار ریاضی معمولاً $\theta$دارای مقدار عددی است و آن را پارامتر و مجموعه‌ی $\Theta$را فضای پارامتر می‌نامند.

اغلب بدون توجه به $\theta$، تصمیم‌گیری می‌کنیم؛ اما عقل سلیم به ما حکم می‌کندکه در صورت امکان نخست بررسی آماری انجام داده تا بتوانیم در مورد $\theta$ آگاهی بیابیم و از این طریق تصمیم‌گیری نمائیم، برای مثال اگر بخواهیم در بهمن ماه به‌گردش برویم، می‌توانیم از بررسی‌آماری اداره هواشناسی برای بهمن ماه استفاده‌کنیم. بعنوان نمونه درتصمیم‌گیری برای‌گردش $\theta$را وضع هوا در نظرمی‌گیریم وفرض می‌کنیم $\Theta = \{\theta_1, \theta_2, \theta_3\}$

به‌گونه‌ای باشدکه $\theta_1$ هوای آفتابی، $\theta_2$ هوای بارانی و $\theta_3$ هوای برفی رانشان دهند. حال با توجه به اینکه وضع هوا ممکن است یکی از سه حالت بیان شده باشد، تصمیم می‌گیریم و برای گردش وسایل مورد نیاز را فراهم می‌کنیم. به طورکلی، نتیجه هر بررسی آماری را که در نظر می‌گیریم با بردار تصادفی $\underline{X} = (X_1,...,X_n)$ نشان می‌دهیم. چگالی توأم این بردار باید به $\theta$ وابستگی داشته باشد تا بتوان درباره‌ی $\theta$ آگاهی یافت. فرض می‌کنیم مقدار مشاهده شده‌ی $\underline{X}$ بردار عددی $\underline{x} = (x_1,...,x_n)$ باشد. مجموعه‌ای را که $\underline{x}$ به آن تعلق دارد با $\chi$نمایش می‌دهیم و آن را فضای‌نمونه‌ای بردار $\underline{X}$ می‌نامیم. مجموعه‌ی $\chi$ زیرمجموعه‌ی $\Re^n$، یعنی فضای $n$بعدی است. اغلب $\underline{X}$ را نمونه‌ای تصادفی در نظر می‌گیرند یعنی فرض می‌کنند $X_1,...,X_n$ مستقل و هم توزیع می‌باشندکه دارای چگالی زیرهستند :

$$X_i \sim f(x,\theta) \qquad \theta \in \Theta \qquad i = 1,...,n$$

به عنوان مثال اگر $\theta$ درصد لامپهای معیوب باشد، برای بررسی آماری $\theta$ می‌توانیم چند بسته ۱۲تایی لامپ را بیازمائیم، تعداد لامپهای معیوب در هر بسته‌ی ۱۲تایی متغیر تصادفی دوجمله‌ای $Bin(12,\theta)$ است. قبل‌ازمعرفی برخی نمادهای نظریه‌ی تصمیم و ادامه‌کار؛ ابتدا برخی اصول برتری و روابط‌حاکم بین‌آنها را با ذکر چندمثال، تعریف و قضیه بیان می‌کنیم.

**مثال ۱-۱-۱ :**

فرض کنید یکی از دو عمل $C_1$ و $C_2$را بتوان انجام داد که دارای پیامدهای شانسی هستند. این پیامدها ممکن است به صورت سود، زیان، نیکوکاری، گرفتاری، شهرت، رسوایی...باشند. در این مثال برای سهولت پیامدها را به صورت سود و زیان و واحد پول را ۱۰۰۰تومان در نظر می‌گیریم. پیامدهای شانسی دو عمل $C_1$ و $C_2$را به ترتیب با متغیرهای $X_1$و $X_2$ بیان می‌کنیم که دارای جدولهای احتمالی زیر هستند:

**جدول ۱-۱**

| $X_1 = x_1$ | $-1$ | $3$ | $4$ |
|---|---|---|---|
| $P(X_1 = x_1)$ | $\frac{1}{2}$ | $\frac{1}{4}$ | $\frac{1}{4}$ |

**جدول ۱-۲**

| $X_2 = x_2$ | $-2$ | $4$ | $9$ |
|---|---|---|---|
| $P(X_2 = x_2)$ | $\frac{1}{2}$ | $\frac{1}{3}$ | $\frac{1}{6}$ |

اینک این پرسش را مطرح می‌کنیم: اگر سود و زیان را به صورتی ملاک قرار دهیم‌که درجدولهای احتمال ارائه شده‌اند، کدام عمل را برمی‌گزینید؟ ازآنجا که: $E(X_1) = 1.25$ و $E(X_2) = 1.83$ پس به طور متوسط $C_2$، بیشتراز $C_1$سود می‌دهد. بنابراین سودجویی ما را به گزینش $C_2$ وادار می‌کند.بعبارت دیگر $C_2$ را براساس امید ریاضی بر $C_1$ ترجیح می‌دهیم. اما ممکن است محافظه‌کار بوده و بخواهیم از زیان شانسی

بپرهیزیم. در اینصورت $C_1$ را بر $C_2$ ترجیح می‌دهیم؛ زیرا در $C_1$، میزان زیان با احتمال پنجاه درصد، ۱ و در $C_2$ با همین احتمال، ۲ است. عددی راکه به $C_1$ و $C_2$ برای تشخیص برتری یکی بر دیگری نسبت‌می‌دهیم مطلوبیت(ارزندگی) $C_1$ و $C_2$ می‌نامیم و با $U(C_1)$ و $U(C_2)$ نشان می‌دهیم. در مثال بالابراساس‌امید ریاضی داریم: $U(C_1) = 1.25$ و $U(C_2) = 1.83$. در همین مثال براساس زیان شانسی داریم: $U(C_1) = -1$ و $U(C_2) = -2$.

**تعریف ۱-۱-۲ :**

برای محاسبه‌ی سود براساس امید ریاضی ممکن است به جای امیدریاضی $X$، امید ریاضی تابعی معقول از $X$ را بکار برد. منطقی است که این تابع را تابعی غیرنزولی از سود و یا زیان برگزینیم. ما این تابع را با $u(x)$ نشان می‌دهیم و آنرا تابع سود می‌نامیم. بنابراین براساس امید ریاضی چنین تابعی سود $C_1$ و $C_2$ را می‌توان به صورت زیر بیان کرد: $U(C_1) = E[u(X_1)]$ و $U(C_2) = E[u(X_2)]$

**مثال ۱-۱-۲ :**

توابع زیر مثالهایی از تابع سود می باشند:

$$u(x) = ax + b \qquad a > 0$$
$$u(x) = x^3$$
$$u(x) = e^x$$
$$u(x) = \log[|x| + 1]$$

در مثال ۱-۱-۱ اگر فرض کنیم $u(x) = x^3$، آنگاه با استفاده از جدولهای احتمال داده شده داریم:

$$U(C_1) = E[u(X_1)] = E[X_1^3] = 22.25$$
$$U(C_2) = E[u(X_2)] = E[X_2^3] = 98.50$$

بنابراین با این تابع سود، $C_2$ بر $C_1$ برتری دارد. در حالتی که محافظه‌کار باشیم و بخواهیم از زیان شانسی بپرهیزیم؛ $u(x)$ را تابعی ثابت برابر بزرگترین زیان در جدول احتمال فرض می‌کنیم. بنابراین در این حالت: $u(X_1) = -1$ و $u(X_2) = -2$ و در نتیجه داریم:

$$U(C_1) = E[u(X_1)] = -1$$
$$U(C_2) = E[u(X_2)] = -2$$

معمولاً $C_1$ یا $C_2$ را در مثال بالا نمود یا پیش‌بینی( Prospect ) می‌نامند، زیرا دورنمایی از هرپیامد همراه با شانس آنرا به ما نشان می‌دهند.

**تعریف ۱-۱-۳ :**

فرض کنید $\underline{S}=(a_1,...,a_k)$ مجموعه پیامدهای ممکن برای یک عمل باشد ( $\underline{S}$ را می توان به منزله‌ی فضای نمونه در یک آزمایش تصادفی یا قرعه‌کشی فرض‌کرد) برای هرتوزیع احتمال $\underline{P}$ روی $\underline{S}$، یعنی‌برای هربردار احتمال $\underline{p}=(p_1,...,p_k)$، یک جدول احتمال بصورت زیر داریم:

**جدول ۱-۳**

| | |
|---|---|
| عمل | $a_1 \ldots a_k$ |
| احتمال | $p_1 \ldots p_k$ |

که $p_i \geq 0$ و $\sum_{i=1}^{k} p_i = 1$ . این جدول احتمال را برای سهولت با: $C=(a_1,...,a_k;p_1,...,p_k)$ و یا $C=(\underline{S};\underline{p})$ نشان می‌دهیم؛ C را یک نمود می‌نامیم. هر نمود به منزله‌ی یک قرعه‌کشی است. مجموعه‌ی تمام نمودهای مربوط به $\underline{S}$ را با D و مجموعه‌ی تمام بردارهای احتمال را با P نشان می‌دهیم. واضح است که D و P تناظری یک به یک دارند. نمود ساده و آمیخته به ترتیب بصورت زیر تعریف می‌شوند:

**تعریف ۱-۱-۴ :**

هر عمل به تنهایی بعنوان یک نمود است زیرا برای مثال $a_1$ را می‌توان بصورت زیر نشان داد: $(a_1,...,a_k;1,0,...,0)$ از اینرو این نمود را با $a_1$ نشان می‌دهیم و آنرا نمود ساده و یا کاری با احتمال یک می‌نامیم.

**تعریف ۱-۱-۵ :**

دو نمود زیر را در نظر می گیریم:

$$C_1=(a_1,...,a_k;p_1,...,p_k)=(\underline{S};\underline{p})$$
$$C_2=(a_1,...,a_k;q_1,...,q_k)=(\underline{S};\underline{q})$$

ترکیبی محدب از بردارهای احتمال این دو نمود، بردار احتمال آمیخته‌ی زیر است:

$$R=\alpha\underline{p}+(1-\alpha)\underline{q} \qquad 0\leq\alpha\leq 1$$

نمود متناظر با این بردار احتمال آمیخته را نمود آمیخته با احتمال آمیختگی $\alpha$ می‌نامیم و آنرا با یکی از نمادهای زیر نشان می‌دهیم:

$$C=(C_1,C_2;\alpha,1-\alpha)$$
$$C=\alpha C_1+(1-\alpha)C_2$$

بطورکلی، می‌توان نشان داد که هر ترکیب محدب از نمودهای $C_1,...,C_k$ نمودی است که آن را نمود آمیخته از آنها می گویند و با $C=(C_1,...,C_k;\alpha_1,...,\alpha_k)$ یا $C=\alpha_1C_1+...+\alpha_kC_k$ نشان می‌دهند. $\alpha_i$ها نامنفی بوده و مجموع آنها برابر یک است.

هنگامی که کارها بصورت سود و زیان نیستند، تشخیص برتری نمودی بر نمود دیگر و تعیین ارزندگی نمودها همواره ممکن و یا برای همه یکسان نیست. در این موارد باید برتری را در چارچوب چند اصل بیان کنیم تا بتوانیم سود و ویژگیهای آنرا به گونه‌ای منطقی با روش اصل گرایی( Axiomatization ) شرح دهیم.

**تعریف ۱-۱-۶ :**

همانگونه که در هندسه دیده‌ایم، این روش مبتنی بر چند مفهوم ابتدایی است که ویژگیهای آنها با چند گزاره بیان می‌شود. درستی این گزاره‌ها که هر یک اصل موضوع( Axiom ) نامیده می‌شود؛ معمولاً از راه تجربه به طور شهودی پذیرفته می‌شود. برای نمونه، در هندسه اقلیدسی، نقطه و خط دو مفهوم ابتدایی و اصل اقلیدس، یک اصل موضوع است.

فرض کنید بردار $\underline{S}$ مجموعه پیامدهای یک عمل و $D$ مجموعه نمودهای مربوط به $\underline{S}$ باشد. اگر $C_1$ و $C_2$ در $D$ باشند رابطه‌ی $C_2 \geq C_1$ را میان آنها تعریف می‌کنیم؛ به معنای اینکه $C_2$ دست کم بر $C_1$ ترجیح دارد. چنین رابطه‌ای را رابطه‌ی برتری یا رابطه‌ی رجحان ( Preference relation ) می‌نامند. اگر داشته باشیم $C_2 \geq C_1$ و $C_1 \geq C_2$ می‌گوییم $C_1$ و $C_2$ بی‌تفاوت ( Indifferent ) و یا معادل ( Equivalent ) هستند و می‌نویسیم: $C_1 \sim C_2$. اگر داشته باشیم $C_2 \geq C_1$ اما نداشته باشیم $C_1 \sim C_2$ می‌گوییم $C_2$ برتر از $C_1$است (یا $C_2$ بر $C_1$ برتری یا رجحان دارد) و می نویسیم: $C_2 > C_1$

دو نماد $\leq$ و $<$ را به روش زیر تعریف می‌کنیم:

$$C_1 \leq C_2 \Leftrightarrow C_2 \geq C_1$$

$$C_1 < C_2 \Leftrightarrow C_2 > C_1$$

**تعریف ۱-۱-۷ :**

برای بیان رابطه‌ی برتری اصول زیر را در نظر می‌گیریم که با عقل سلیم سازگارند:

اصل الف- برای هر $C_1$ و $C_2$فقط یکی از سه رابطه‌ی زیر برقرار است:

$$(۱) C_1 > C_2$$

$$(۲) C_1 < C_2$$

$$(۳) C_1 \sim C_2$$

این اصل را، اصل یک از سه می گویند ( Trichotomy )

اصل ب- برای هر $C_1$و $C_2$و $C_3$ اگر $C_1 \geq C_2$ و $C_2 \geq C_3$ آنگاه:

$$C_1 \geq C_3$$

این اصل را اصل انتقال پذیری می‌نامند. ( Transitivity )

اصل ج- برای هر $C_1$ و $C_2$ و $C$ با فرض $C_2 > C_1$ و برای هر عدد $0 < \alpha < 1$ داریم:

$$(C_2, C; \alpha) > (C_1, C; \alpha)$$

از این اصل نتیجه می‌شود که:

$$(C, C_2; \alpha) > (C, C_1; \alpha)$$

این اصل را اصل آمیختگی می‌نامند. ( Mixture )

اصل د- اگر $C_2 > C > C_1$ آنگاه یک عدد $0 < \alpha < 1$ یافت می‌شود به گونه‌ای که:

$$C \sim (C_2, C_1; \alpha)$$

این اصل را اصل پیوستگی می‌نامند. ( Continuity )

اصل هـ -

| | |
|---|---|
| (بازتابی) | $C \sim C$ |
| (تقارن) | $C_1 \sim C_2 \Leftrightarrow C_2 \sim C_1$ |
| (انتقال پذیری) | $C_1 \sim C_2 \ , \ C_2 \sim C_3 \Rightarrow C_1 \sim C_3$ |
| (انتقال پذیری) | $C_1 > C_2 \ , \ C_2 > C_3 \Rightarrow C_1 > C_3$ |
| (معادل گذاری) | $C_1 \geq C_2 \ , \ C_2 \sim C_3 \Rightarrow C_1 \geq C_3$ |
| (معادل گذاری) | $C_1 > C_2 \ , \ C_2 \sim C_3 \Rightarrow C_1 > C_3$ |

فرض $D$ مجموعه‌ی تمامی اعمالی باشد که بتوانیم یکی از آنها را انجام دهیم، مجموعه‌ی $D$ را فضای اعمال (کارها) می گوییم. با مشاهده‌ی $\underline{x}$ بنابر ضابطه‌ای به عمل $\delta \in D$ می‌پردازیم. ضابطه‌ای که براساس آن مشاهدات $\underline{x}$ به عملی می‌انجامند، درحقیقت قانون تصمیم گیری را مشخص می کند وتابع تصمیم نامیده می‌شود.

**تعریف ۱-۱-۸ :**

تابع تصمیم؛ تابعی با دامنه‌ی $\chi$ (فضای مشاهدات ) و برد $D$ (فضای اعمال ) است این تابع را با $\delta(\underline{x})$ نشان می‌دهیم، بنابراین برای هر $\underline{x} \in \chi$ به یک عمل $\delta$ می‌پردازیم یعنی داریم:

$$\delta(\underline{x}) = \delta \in D$$

از نظر آمار ریاضی $\delta(\underline{x})$ مقدار مشاهده شده‌ی آماره‌ی $\delta(\underline{X})$ است. هنگامی که $\theta$ پارامتر یک توزیع و $\delta(\underline{x})$ برآورد آن باشد عمل $\delta$، انتخاب $\delta(\underline{x})$ به عنوان برآورد $\theta$ می‌باشد. مجموعه‌ی تمام توابع تصمیم را با $D$ نشان می‌دهیم و آن را فضای توابع تصمیم می گوییم.

هر عمل و یا تصمیم به پیامدی می‌انجامد که ارزیابی آن با مقیاس صحیح همواره ساده یا ممکن نیست. بعنوان مثال درباره‌ی نتیجه‌ی کار معلم، قاضی، سیاستمدار و ... نمی‌توان به سادگی داوری کرد بویژه اگر شانس نیز در این کارها مؤثر باشد، آنگاه سنجش آن دشوارتر می‌شود. در مواردی که نتیجه‌ی عملی با پول ارزیابی شده ظاهراً مشکل کمتر می‌شود، اما از آنجائی که پول برای همه و همیشه ارزش ثابتی ندارد بنابراین همواره مقیاس معقولی نیست. برای مثال در مورد شخصی بادرآمدکم، شاید بدست آوردن ۵۰۰۰ تومان در روز با ارزش باشد، اما اگر همین شخص از راهی مشروع و یا نامشروع یک میلیون تومان بدست آورد، ممکن است ۵۰۰۰تومان در روز را ناچیز پندارد. شخص دیگری ممکن است ازدرآمد کلان چشم پوشی‌کرده و هرکاری را نپذیرد تا آبروی اجتماعی خود را حفظ کند. بنابراین مسأله برتری عملی بر عمل دیگر به عوامل گوناگونی وابسته است. عددی را که به پیامد کاری می‌توان نسبت داد، باتوجه به ارزش و اهمیت این پیامد برای‌اعمالی که شانس درآنها دخیل می‌باشد، مطلوبیت(ارزندگی) می‌نامیم. این عدد همیشه و برای همه‌ی ما یکسان نیست اما به هرحال، برای تعیین‌آن می‌بایست از اصولی منطقی و معقول استفاده کرد. هنگامی‌که با مشاهده‌ی $\underline{x}$ برای انجام عمل $\delta$ تصمیم می‌گیریم ممکن است با زیان مالی یا غیر مالی؛ برای مثال رسوائی یا نگرانی روبرو شویم؛ زیرا از وضع طبیعت‌آگاهی نداریم این زیان را با یک تابع دو متغیره‌ی نامنفی از $\theta$ و $\delta = \delta(\underline{x})$ بیان می‌کنیم. چنین تابعی را که می‌تواند به صورتهای گوناگون باشد تابع زیان می‌نامند. زیان، همانگونه که گفتیم ممکن است مالی نباشد، بعنوان مثال می‌تواند به صورت رسوائی، نگرانی، ناراحتی و سرگردانی و ... باشد، برای اینکه بتوان براساس آن تصمیم‌گیری کرد باید دست‌کم با یک مقیاس ترتیبی قابل اندازه‌گیری باشد. معمولاً برای تصمیم‌گیری دربرخی رشته‌ها مانند اقتصاد مطلوبیت را بکار می‌برندکه معیاری برای سود است. اما آماردانان در اثرمحافظه‌کاری‌ترجیح می‌دهندکه‌برحسب تابع زیان تصمیم‌گیری کنند.

**تعریف ۱-۱-۹ :**

تابع زیان؛ تابع دو متغیره‌ی‌کراندار و نامنفی با دامنه‌ی $\Theta \times D$ (حاصلضرب دکارتی فضای پارامتر و فضای اعمال) و برد $\Re^+$ است که آنرا با نماد $L(\theta, \delta)$ نشان می‌دهند. درصورتی که بتوانیم یک رابطه‌ی برتری‌که تابع اصول برتری باشد، روی $(\theta, \delta)$ها تعریف‌کنیم، آنگاه مطلوبیت $(\theta, \delta)$ عدد $U(\theta, \delta)$است. اکنون زیان را بصورت مطلوبیت با علامت منفی در نظر می‌گیریم؛ یعنی: $L(\theta, \delta) = -U(\theta, \delta)$

طبیعی است که اگر دو نفر درباره‌ی امری دو تابع مطلوبیت متفاوت داشته باشند، ممکن است به تصمیم‌گیریهای متفاوتی بپردازند در حالیکه هردو خردمندانه و معقول تصمیم‌گیری می‌کنند.

هرچه $\delta$و $\theta$ به هم نزدیکتر باشند زیان ناشی از انجام عمل $\delta$ کمترخواهد بود. البته تابع زیانی مناسب باید این امر را نشان دهد. برای مثال در توزیع دوجمله ای؛ $\theta$را احتمال پیروزی فرض می‌کنیم و داریم:

$\Theta = D = (0,1)$

اینک تابع زیان $L(\theta,\delta) = |\delta - \theta|$ را در نظر می گیریم واضح است که این تابع، منظور ما را برآورده می کند. چنین توابعی زیاد هستند، اکنون برخی ازآنها را معرفی و مطالعه می کنیم.

الف- تابع زیان درجه‌ی دو :

تابع زیان $L(\theta,\delta) = (\delta - \theta)^2$ را تابع زیان درجه‌ی دوم یا توان دوم خطا می‌نامند. این تابع بـرای بـرآوردیـابی کاربرد فراوانی دارد، با اینکه به عنوان تابع زیان ممکن است کراندار نباشـد. برخـی از دلایلـی‌کـه تـابع زیـان درجه‌ی دو در برآوردیابی ونظریه‌ی تصمیم، بسیار مورد استفاده قرار می گیرد عبارتند از:

۱- در مسائل برآوردیابی زمانی‌که برآوردگرهای نا اُریب پارامتر $\theta$ تحت بررسی و مورد مطالعه قرار گرفتند چون:

$$R(\delta,\theta) = E\left[L(\theta,\delta(\underline{X})\right] = E_\theta\left[\delta(\underline{X}) - \theta\right]^2$$

آماردانان مایل بودندکه این تابع ریسک همان واریانس برآوردگر باشد.

۲- دومین دلیل به ارتباط بین این تابع زیان با تئوری حداقل مربعات بر می گردد. شباهت ایـن دو، تـابع زیـان مربع خطا را در نظر آماردانان‌ها معقول‌تر می‌سازد.

۳- دراکثر مسائل نظریه‌ی تصمیم استفاده از زیان درجه‌ی دو منجر به سادگی وآسان بودن محاسبات مربوطـه می گردد و شاید مهمترین دلیل استفاده‌ی فراوان از تابع زیان درجه‌ی دو همین باشد.

تابع زیان درجه‌ی دو را می توان به صورت زیر تعمیم داد:

$$L(\theta,\delta) = \omega(\theta)(\delta - \theta)^2$$

این تابع زیان را، تابع زیان درجـه‌ی دو وزنـی مـی‌نامنـد. وزن $\omega(\theta) > 0$ تـابعی از پـارامتر $\theta$ مـی‌باشـد. وزن $\omega(\theta)$ در واقع این حقیقت را نشان می‌دهدکه یک خطای معلوم در مسأله‌ی برآورد، اغلب زیان را طبق آنچه در آینده برای $\theta$رخ خواهد داد؛ منعکس می کند. برای مثال، تـابع زیـان زیـر $L(\theta,\delta) = \frac{(\delta - \theta)^2}{|\theta| + 1}$ یـک تـابع زیان درجه دو وزنی با وزن :

$$\omega(\theta) = \frac{1}{|\theta| + 1}$$

است. بویژه برای بردارهای $\underline{\theta} = (\theta_1, \ldots, \theta_K)$ و $\underline{\delta} = (\delta_1, \ldots, \delta_K)$ و $\omega(\theta) = (\omega_1(\theta), \ldots, \omega_K(\theta))$ تابع:

$$L(\theta,\delta) = \sum_{i=1}^{k} \omega_i(\theta)(\delta_i - \theta_i)^2$$

را بعنوان تابع زیان وزنی درجه‌ی دو بکار می‌برند.

ب- تابع زیان خطی:

تابع:

$$L(\theta,\delta)=\begin{cases} c(\theta-\delta) & \theta\ge\delta \\ d(\delta-\theta) & \theta<\delta \end{cases}$$

که: $c,d>0$؛ را تابع زیان خطی می‌نامند. دو عدد ثابت c و d به گونه‌ای انتخاب می‌شوند تا بیش برآوردی ( Over Estimation ) و کم برآوردی ( Under Estimation ) برحسب اهمیت آنها در تابع زیان منعکس شوند. اگر $c=d=1$، تابع زیان خطی به صورت $L(\theta,\delta)=|\delta-\theta|$ در می‌آید که تابع زیان قدرمطلق خطا نامیده می‌شود.

ج- تابع زیان صفر و یک:

فرض کنید $D=\{\delta_0,\delta_1\}$ فضای کارها باشد. بعنوان مثال برای آزمون $H_0:\theta=\theta_0$ در برابر $H_1:\theta=\theta_1$، $\Theta=\{\theta_0,\theta_1\}$ فضای پارامتر است، عمل $\delta_0$ را پذیرفتن فرض $H_0$ و عمل $\delta_1$ را پذیرفتن فرض $H_1$ در نظر می‌گیریم. تابع زیر را که برای $\delta_0$ و $\delta_1$ به صورت زیر تعریف شده است تابع زیان صفر و یک می نامند:

$$L(\theta,\delta_0)=\begin{cases} 0 & \theta=\theta_0 \\ 1 & \theta=\theta_1 \end{cases}$$

و

$$L(\theta,\delta_1)=\begin{cases} 1 & \theta=\theta_0 \\ 0 & \theta=\theta_1 \end{cases}$$

با این تابع اگر عمل درست را انجام دهیم زیان صفر است و اگر عمل نادرست را انجام دهیم، زیان یک است. برای مثال اگر $\delta_0$ را انجام دهیم؛ یعنی بپذیریم که $\theta=\theta_0$، اما درحقیقت $\theta=\theta_1$ باشد؛آنگاه‌زیان یک است. تابع زیانی که از تابع بالا بهتر بوده، بصورت زیر است که در آن $l_0$ و $l_1$ دو عدد مثبت هستند که بزرگی آنها به اهمیت زیان عمل نادرست بستگی دارد:

$$L(\theta,\delta_0)=\begin{cases} 0 & \theta=\theta_0 \\ l_0 & \theta=\theta_1 \end{cases}$$

$$L(\theta,\delta_1)=\begin{cases} l_1 & \theta=\theta_0 \\ 0 & \theta=\theta_1 \end{cases}$$

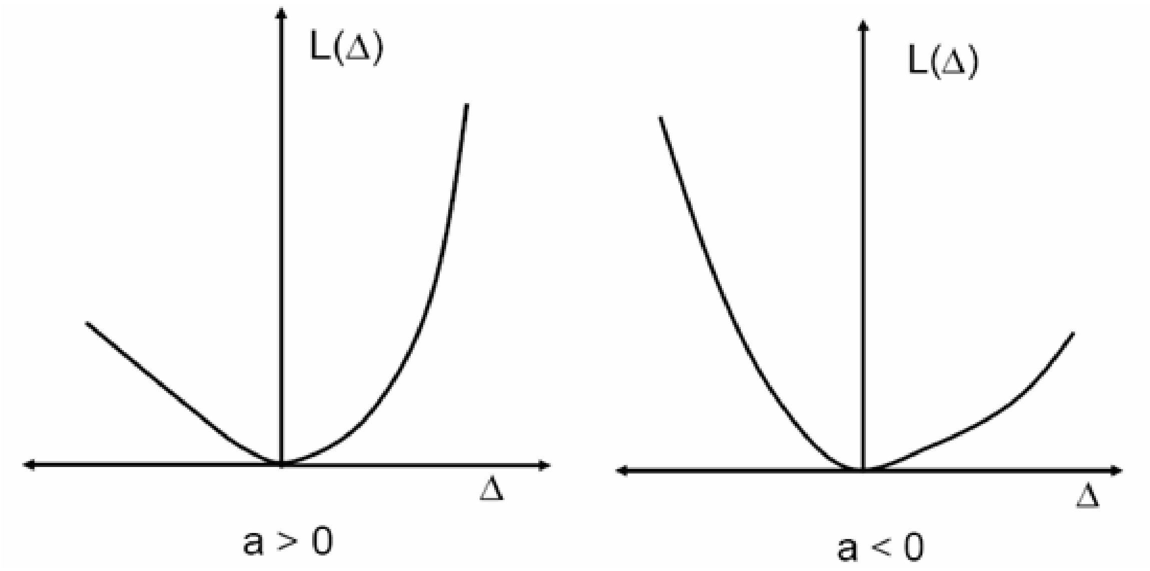
د- تابع زیان لاینکس یا خطی نمائی ( linex ):

این تابع زیان به صورت زیر است:

$$L(\theta,\delta)=b[e^{a(\delta-\theta)}-a(\delta-\theta)-1]$$

در این تابع $b>0$، پارامتر مقیاس و $a\neq 0$ پارامتر شکل است. از آنجا که در این تابع هم شکل نمائی و هم شکل خطی هر دو وجود دارند، از اینرو آنرا تابع زیان لاینکس ( linex ) به معنی خطی نمائی می نامند: Linex:Linear Exponential این تابع نسبت به a محدب اکید است. زیان وارد با این تابع برای

بیش‌برآوردی محسوستر از کم‌برآوردی است، یعنی تابع زیان نسبت به این دو امرنامتقارن است. شاید در برخی موارد بر تابع زیان درجه‌ی دو و تابع زیان قدرمطلق‌خطا برتری داشته باشد که هر دو متقارن هستند.
اگر تعریف کنیم $\Delta = \delta - \theta$، آنگاه فرم کلی این تابع زیان برای $a \neq 0$ به صورت زیر است:




**شکل ۱-۱**

که $L(\Delta)$ عبارتست از:

$$L(\Delta) = b[e^{a\Delta} - a\Delta - 1]$$

فرض می‌کنیم یک تابع زیان مناسب برای مساله برآورد تعریف‌کرده‌ایم، و تابع زیان را بعنوان یک اندازه‌ی خطا یا زیان در نظر می‌گیریم. هدف ما انتخاب برآوردگر $\delta = \delta(\underline{X})$ است که این زیان را مینیمم سازد. تابع زیان خودش به برآورد $\delta(\underline{x})$ بستگی دارد و $\delta(\underline{x})$ یک مقدار برآوردگر $\delta(\underline{X})$؛ یعنی: $\delta = \delta(\underline{X})$ می‌باشد. بنابراین زیان به نمونه‌ی $X_1,...,X_n$ بستگی دارد. ما نمی‌توانیم امیدوار باشیم که زیان برای هر نمونه‌ی ممکن مینیمم باشد، اما می‌توانیم سعی کنیم که به طور متوسط زیان را کاهش دهیم. بنابراین اگر هدفمان را از انتخاب برآوردگری که زیان را مینیمم می‌سازد به انتخاب برآوردگری که متوسط زیان را مینیمم می‌کند تغییر دهیم، می‌توانیم وابستگی زیان را به نمونه‌ی $X_1,...,X_n$ از بین ببریم. برای نیل به این هدف؛ تابع تصمیم $\delta$ و تابع زیان $L(\theta,\delta)$ را در نظر می‌گیریم که هر دو از قبل مشخص شده‌اند. اگر با مشاهده‌ی $\underline{x}$؛ تصمیم به عمل $\delta = \delta(\underline{x})$ بگیریم و وضع طبیعت $\theta$ باشد، مقدار زیان عدد $L(\theta,\delta(\underline{x}))$ است. با تبدیل $\underline{x}$ به $\underline{X}$، این عدد به متغیر تصادفی $L(\theta,\delta(\underline{X}))$ تبدیل می‌شود. در حقیقت پیش از انجام آزمایش، زیان بصورت یک متغیر تصادفی است. بنابراین باید امید ریاضی این متغیرتصادفی را محاسبه کرد، تا زیان به طور متوسط بدست

آید. واضح است که این امید ریاضی فقط به $\theta$ بستگی دارد، یعنی تابعی از $\theta$ است چنین تابعی را تابع ریسک ( Risk function ) و یا تابع مخاطره می‌نامیم. تابع ریسک بصورت زیر تعریف می‌شود:

**تعریف ۱-۱-۱۰ :**

برای هر تصمیم مشخص $\delta$ تابع ریسک عبارتست از تابعی نامنفی و متناهی با دامنه‌ی $\Theta$ و برد $\Re$ که آنرا با $R(\delta,\theta)$ نشان می‌دهیم و بصورت زیر تعریف می‌کنیم:

$$R(\delta,\theta) = E_{\underline{X}}[L(\theta,\delta(\underline{X}))]$$

چگالی مربوط به امید ریاضی بالا، چگالی بردار $\underline{X}$ و یا چگالی‌آماره‌ی $\delta(\underline{X})$ بوده که معمولاً به پارامتر مجهول $\theta$ بستگی دارد. می‌بایست توجه کنیم که تابع $R(\delta,\theta)$ به تابع زیان و تابع تصمیم بستگی دارد و با تغییر هرکدام تغییر می‌کند.

فرض کنید $X_1,...,X_n$ یک نمونه‌ی تصادفی از چگالی $f(x|\theta)$ با $\theta \in \Theta$، باشد که در آن تابع $f(.|\theta)$ به استثنای $\theta$، معلوم فرض می‌شود. فرض می‌کنیم‌که $\theta$ی نامعلوم مقداری از یک متغیر تصادفی $\Theta$ است که توزیع $\Theta$ معلوم و شامل هیچ پارامتر مجهولی نیست. امیدواریم بر اساس نمونه‌ی تصادفی $X_1,...,X_n$ بتوانیم $\gamma(\theta)$ را که تابع معینی از $\theta$ است، برآوردکنیم. بعلاوه فرض می‌کنیم که تابع زیان $L(\theta,\delta)$ مشخص شده است، که در آن $L(\theta,\delta)$ زیانی را که در صورت برآورد $\gamma(\theta)$ با $\delta$، وقتی $\theta$ برابر پارامتر چگالی‌ای باشد که از آن نمونه گرفته‌ایم، متحمل می‌شویم، نشان می‌دهد. دیدیم که برای هر برآوردگر $\delta = \delta(\underline{X})$، $E_{\underline{X}}[L(\theta,\delta(\underline{X}))]$ متوسط زیان آن برآوردگر را نشان می‌دهد و این متوسط زیان را ریسک برآوردگر $\delta(\underline{x})$، که با $R(\theta,\delta)$ نشان داده شد، تعریف کردیم. به علاوه دریافتیم که دو برآوردگر مانند $\delta_1$ و $\delta_2$ را می‌توان با در نظر گرفتن ریسک‌های مربوط به آنها یعنی $R(\theta,\delta_1)$ و $R(\theta,\delta_2)$ مقایسه کرد و برتری را به برآوردگری با ریسک کمتر داد. به طور کلی، توابع ریسک دو برآوردگر به عنوان توابعی از $\theta$ می‌توانند یکدیگر را قطع کنند. یک تابع ریسک برای برخی $\theta$ها کوچکتر است و برای برخی دیگر از $\theta$ها تابع ریسک دیگر کوچکتر است. پس، چون $\theta$ نامعلوم است، انتخاب بین این دو برآوردگر مشکل است. این مشکل به دلیل بستگی تابع ریسک به $\theta$ است. حال، چون فرض کرده‌ایم که $\theta$ مقداری از یک متغیر تصادفی $\Theta$ است و توزیع آن نیز معلوم فرض شده است، یک راه حل طبیعی رفع وابستگی تابع ریسک به $\theta$ عبارتست از محاسبه‌ی میانگین نسبت به $\theta$، که با استفاده از چگالی $\theta$ صورت می‌گیرد.