چکیده

در این پایان‌نامه مسئله برآورد پارامترهای رگرسیونی در مدل رگرسیون خطی چندگانه هنگامی که هم‌خطی چندگانه موجود باشد، بررسی می‌شود. تحت فرض نرمال، سه برآوردگر بیز تجربی شامل برآوردگر کمترین توان‌های دوم انقباضی، برآوردگر بیز تجربی سلسله مراتبی انقباضی و برآوردگر بیز تجربی تجزیه شده ارائه می‌شوند. برآوردگرها به ترتیب براساس روش مولفه‌های اصلی، برآوردگر کمترین توان‌های دوم دوگانه و با انتخاب پیشین‌های متفاوت برای پارامترهای مدل رگرسیونی به دست می‌آیند.

تاکنون ثابت نشده است که برآوردگرهای پیشنهاد شده از نظر همگرایی بهتر از برآوردگرهای کمترین توان‌های دوم هستند که در این پایان نامه این موضوع ثابت می‌شود. همچنین نشان داده می‌شود که این سه برآوردگر هنگامی که هم‌خطی چندگانه موجود باشد و از طریق شبیه‌سازی و مطالعات تجربی به دست آیند از نظر کارایی برآوردگرهای مفیدی هستند.

واژه‌های کلیدی: رگرسیون چندگانه، هم‌خطی چندگانه، رگرسیون ستیغی، روش بیز تجربی، روش مولفه‌های اصلی، مینیماکس.

فهرست مندرجات

لیست اشکال

فصل ۱

تعاریف و پیش‌نیازها

۱.۱ مولفه‌های اصلی جامعه

اگر بردار تصادفی $\boldsymbol{X}' = [X_1, X_2, \ldots, X_p]$ دارای ماتریس کوواریانس $\boldsymbol{\Sigma}$ و مقادیر ویژه‌ی $\lambda_1 \geq \lambda_2 \geq \ldots, \lambda_p \geq 0$ باشد. با در نظر گرفتن ترکیبات خطی زیر:

$$
\begin{aligned}
Y_1 &= l_1'X = l_{11}X_1 + l_{21}X_2 + \ldots + l_{p1}X_p \\
Y_2 &= l_2'X = l_{12}X_1 + l_{22}X_2 + \ldots + l_{p2}X_p \\
&\vdots \\
Y_i &= l_i'X = l_{1i}X_1 + l_{2i}X_2 + \ldots + l_{pi}X_p \\
&\vdots \\
Y_p &= l_p'X = l_{1p}X_1 + l_{2p}X_2 + \ldots + l_{pp}X_p
\end{aligned}
$$

داریم:

$$Var(Y_i) = l_i' \boldsymbol{\Sigma}\, l_i \qquad i = ۱, \ldots, p$$

و

$$Cov(Y_i, Y_k) = l_i' \boldsymbol{\Sigma}\, l_k \qquad i, k = ۱, \ldots, p$$

مولفه‌های اصلی $X_p, \ldots, X_۲, X_۱$، آن ترکیبات خطی ناهمبسته‌ی $Y_p, \ldots, Y_۲, Y_۱$ هستند که واریانس آن‌ها بیشترین مقدار ممکن را اتخاذ کند.

قضیه ۱—۱: اگر $\boldsymbol{\Sigma}$ ماتریس کوواریانس بردار تصادفی $\boldsymbol{X}' = [X_۱, X_۲, \ldots, X_p]$ باشد. همچنین اگر $\boldsymbol{\Sigma}$ دارای p زوج مقدار ویژه— بردار ویژه‌ی $(\lambda_p, e_p), \ldots, (\lambda_۲, e_۲), (\lambda_۱, e_۱)$ باشد به طوری که $\lambda_۱ \geq \lambda_۲ \geq \ldots, \lambda_p \geq ۰$ و اگر مولفه اصلی i ام به صورت:

$$Y_i = e_i' \boldsymbol{X} \;= e_{۱i} X_۱ + e_{۲i} X_۲ + \ldots + e_{pi} X_p \qquad i = ۱, \ldots, p$$

تعریف شود، آن‌گاه واریانس i امین مولفه اصلی ماکسیمم خواهد شد. با این انتخاب‌ها داریم:

$$\begin{aligned} Var(Y_i) &= e_i' \boldsymbol{\Sigma}\, e_i \qquad i = ۱, \ldots, p \\ Cov(Y_i, Y_k) &= e_i' \boldsymbol{\Sigma}\, e_k \qquad i \neq k \end{aligned}$$

در صورتی که بعضی از λ_i ها برابر باشند، انتخاب e_i ها و در نتیجه Y_i ها یکتا نخواهند بود (الکساندر مود (۱۹۲۱)).

۲.۱ برآوردگرهای بیز و مینیماکس

می‌دانیم بهترین برآوردگر، برآوردگری است که به طور یکنواخت دارای کمترین مقدار مخاطره باشد. اما متأسفانه موقعیتی که تحت آن بهترین برآوردگر وجود داشته باشد، پیش

نمی‌آید. یعنی برای یک مقدار ثابت θ ، برآوردگری با کمترین مخاطره را می‌توان یافت، اما این بهترین برآوردگر برای مقادیر مختلف θ تغییر می‌کند. بنابراین هیچ برآوردگری را نمی‌توان به عنوان بهترین برآوردگر ارائه داد و به همین دلیل در آمار به یافتن برآوردگرهای بهینه اکتفا می‌شود. برای این منظور در آمار دو روش متداول است:

۱— قرار دادن محدودیت‌های مناسب بر روی برآوردگرها

۲— مرتب کردن برآوردگرها تحت یک قانون معین

در این فصل به اختصار روش دوم را مورد بررسی قرار می‌دهیم. بدین معنی که با اعمال یک قانون خاص، ابتدا برآوردگرها را مرتب کرده سپس برآوردگر مطلوب را انتخاب می‌کنیم. اصول مینیماکس و بیز دو اصل معمول از این روش‌اند.

۳.۱ اصول مینیماکس و بیز

در برآورد پارامتر $\gamma(\theta)$ براساس برآوردگر $\delta(X)$ ، دقت برآوردگر با تابع مخاطره:

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

اندازه گرفته می‌شود. در مسایل برآوردیابی علاقمند به دستیابی به برآوردگری مانند δ هستیم که برای هر $\theta \in \Theta$ ، $R(\theta, \delta)$ را مینیمم کند. می‌دانیم که در حالت کلی، مسئله دارای جواب نیست. نتیجه منطقی این موضوع، پیشنهاد ایجاد محدودیت در کلاس برآوردگرها برای دستیابی به برآوردگرهای بهینه است. گفتیم که یک روش برای پیدا کردن برآوردگرهای بهینه، برقراری یک نوع رابطه ترتیبی بین برآوردگرهاست. دو روش اساسی در برقراری رابطه‌ای ترتیبی بین برآوردگرها، اصول مینیماکس و بیز است.

یک رابطه ترتیبی بین برآوردگرها می‌تواند براساس رخداد بدترین حالت ممکن برای آماردان

در نظر گرفته شود. به عبارت دیگر برآوردگر δ_1 به برآوردگر δ_2 ترجیح داده می‌شود اگر

$$\sup_\theta R(\theta, \delta_1) \leq \sup_\theta R(\theta, \delta_2)$$

بر پایه این موضوع، اصل مینیماکس، برآوردگری از δ را بررسی می‌کند که مقدار $\sup_\theta R(\theta, \delta)$ را مینیمم کند. چنین انتخابی از δ منجر به انتخاب برآوردگر مینیماکس δ_m می‌شود، یعنی برای هر $\delta \in D$ که D فضای کل برآوردگرهاست داشته باشیم:

$$\sup_\theta R(\theta, \delta_m) \leq \sup_\theta R(\theta, \delta)$$

به عبارت دیگر، براساس این اصل، بدترین حالات ممکن از برآوردگرها به ازای $\theta \in \Theta$ را در نظر گرفته و بین بدترین حالات ممکن، آن δ یی اختیار می‌شود که کمترین مقدار مخاطره را اختیار کند. واضح است که در این حالت لزومی ندارد برآوردگر مینیماکس یکتا باشد.

اصل بیز، برآوردگری از θ را بررسی می‌کند که برای یک تابع وزنی نظیر $G(\theta)$ ، مقدار $\int_\Theta R(\theta, \delta) dG(\theta)$ حداقل شود. یعنی براساس تابع وزنی G و تابع زیان L، کلیه مقادیر $\int_\Theta R(\theta, \delta) dG(\theta)$ را برای هر $\delta \in D$ ، از کوچک به بزرگ مرتب و برآوردگر δ—ای را انتخاب می‌کند که مقدار این انتگرال برای آن از همه کمتر باشد که به انتخاب برآوردگر بیز $\delta_B(X)$ می‌انجامد، یعنی برای هر برآوردگر $\delta \in D$،

$$\int_\Theta R(\theta, \delta_B) dG(\theta) \leq \int_\Theta R(\theta, \delta) dG(\theta).$$

۴.۱ توزیع پیشین و توزیع پسین

توزیع پارامتر قبل از مشاهده هر داده‌ای، توزیع پیشین آن پارامتر نامیده می‌شود. توزیع شرطی پارامتر به شرط مشاهده داده‌ها توزیع پسین نامیده می‌شود. اگر مقادیر مشاهده

شده داده‌ها را در تابع احتمال شرطی یا تابع چگالی احتمال شرطی داده‌ها به شرط پارامتر، جایگذاری کنیم نتیجه این کار فقط تابعی از پارامتر است که آن را تابع درستنمایی می‌نامیم.

با توجه به ماهیت اصل بیز، در مسائل استنباط آماری به روش بیز براساس مشاهداتی که از خانواده توزیع‌ها اختیار می‌شود پارامتر θ دارای یک مقدار نامعلوم است. به عبارت دیگر، در حقیقت θ به عنوان یک مقدار متغیر تصادفی W در نظر گرفته می‌شود که مقادیر ممکن آن فضای پارامتر Θ است و دارای توزیع $G(\theta)$ یا تابع احتمال (تابع چگالی احتمال) $g(\theta)$ است و از آن با عنوان توزیع پیشین یا تابع احتمال پیشین (تابع چگالی احتمال پیشین) یاد می‌کنیم. در واقع، توزیع پیشین تبلور کاربر آمار از خلاصه اطلاعات و دانسته‌های او در این باره است که احتمال قرار داشتن θ در چه بخش‌هایی از Θ بیش از همه است. به بیان دیگر قبل از مشاهده و جمع‌آوری هر گونه داده‌ای، اطلاعات و دانسته‌های قبلی کاربر آمار، وی را متقاعد می‌کند بر این باور باشد که شانس قرار گرفتن مقادیر θ در فضای Θ چگونه است و چنین باورهایی را می‌توان در قالب یک تابع توزیع بیان کرد.

تابع مخاطره بیزی در انتخاب δ به عنوان یک برآوردگر نسبت به توزیع پیشین G و تابع زیان L، که آن را با نماد $r(G,\delta)$ نمایش می‌دهیم، برابر است با:

$$\begin{aligned} r(G,\delta) &= E\{E[L(W,\delta(X))]\} \\ &= E[R(W,\delta)] \\ &= \int_{\Theta} R(\theta,\delta)dG(\theta) \end{aligned} \qquad (۱.۱)$$

و به دنبال برآوردگر δ_B هستیم که مقدار (۱.۱) را مینیمم کند. در این حالت از برآوردگر $\delta_B(X)$ با عنوان برآوردگر بیز $\gamma(\theta)$ نسبت به توزیع پیشین G تحت تابع زیان L یاد می‌کنیم.

فرض می‌کنیم که مدل احتمالی آزمایش بستگی به پارامتر θ دارد که از خانواده توزیع‌های $\{F_\theta : \theta \in \Theta\}$ با خانواده چگالی‌های $\{f_\theta : \theta \in \Theta\}$ است، همچنین فرض کنیم که θ خود مقدار مشاهده شده یک متغیر تصادفی، مثلاً W با توزیع معلوم $G(\theta)$ یا چگالی معلوم $g(\theta)$ است، که از آن با نام توزیع پیشین W یاد می‌کنیم در این حالت، $f_\theta(.)$ را می‌توان به عنوان

چگالی شرطی متغیر X به شرط $W = \theta$ تلقی کرده و چگالی توأم X و W را از رابطه زیر به دست آورد.

$$f_{X,W}(x,\theta) = f_{X|W=\theta}(x)g(\theta) \qquad (۲.۱)$$

توزیع شرطی W به شرط $X = x$ قابل محاسبه است، که از آن با نام توزیع پسین W یاد می‌کنیم. توزیع پسین W دارای چگالی پسین است که با استفاده از رابطه (۲.۱) و قضیه بیز به صورت زیر قابل محاسبه است.

$$g_{W|X=x}(\theta) = \frac{f_{X,W}(x,\theta)}{f_X(x)}$$

که در آن $f_{X,W}(.,.)$ تابع چگالی احتمال توأم X و W و $f_X(x)$ تابع چگالی احتمال کناری X است، به طوری که

$$f_X(x) = \begin{cases} \sum_\Theta f_{X,W}(x,\theta) & W \text{ گسسته باشد} \\ \int_\Theta f_{X,\theta}(x,\theta)d\theta & W \text{ پیوسته باشد} \end{cases}$$

بنابراین تابع چگالی پسین را می‌توان به صورت زیر بازنویسی کرد:

$$\begin{aligned} g_{W|X=x}(\theta) &\equiv g(\theta|x) \\ &= \frac{f_\theta(x)g(\theta)}{\int_\Theta f_\theta(x)dG(\theta)}. \end{aligned}$$

۵.۱ برآوردگرهای بیز

فرض کنید $\boldsymbol{X} = (X_1, \ldots, X_n)$ نمایانگر یک نمونه تصادفی n تایی از خانواده چگالی‌های $\{f_\theta : \theta \in \Theta\}$ باشد. اگر توزیع پیشین انتخابی $G(\theta)$ با چگالی $g(\theta)$ و تابع زیان

$L(\theta, \delta)$ باشد، آن‌گاه تابع مخاطره برآوردگر δ در برآورد پارامتر $\gamma(\theta)$ به صورت زیر خواهد بود:

$$\begin{aligned} R(\theta, \delta) &= E[L(\theta, \delta(\boldsymbol{X}))] \\ &= \int_{\chi} L(\theta, \delta(\boldsymbol{x}))f_{\theta}(\boldsymbol{x})d\boldsymbol{x} \end{aligned}$$

و تابع مخاطره بیزی برآوردگر δ در برآورد $\gamma(\theta)$ نسبت به توزیع پیشین G و تابع زیان L، به صورت زیر است (توجه کنید که θ نقش یک متغیر را دارد).

$$\begin{aligned} r(G, \delta) &= E[R(\theta, \delta)] \\ &= \int_{\Theta} R(\theta, \delta)dG(\theta) \\ &= \int_{\Theta} R(\theta, \delta)g(\theta)d\theta \end{aligned}$$

بنابراین به سادگی معلوم می‌شود که:

$$\begin{aligned} r(G, \delta) &= \int_{\Theta} R(\theta, \delta)g(\theta)d\theta \\ &= \int_{\Theta} \{\int_{\chi} L(\theta, \delta(\boldsymbol{x}))f_{\theta}(\boldsymbol{x})d\boldsymbol{x}\}g(\theta)d\theta \\ &= \int_{\Theta}\int_{\chi} L(\theta, \delta(\boldsymbol{x}))f_{\theta}(\boldsymbol{x})g(\theta)d\boldsymbol{x}\, d\theta \\ &= \int_{\Theta}\int_{\chi} L(\theta, \delta(\boldsymbol{x}))f(\boldsymbol{x}, \theta)d\boldsymbol{x}\, d\theta \\ &= \int_{\Theta}\int_{\chi} L(\theta, \delta(\boldsymbol{x}))g(\theta|\boldsymbol{x})f_{\boldsymbol{X}}(\boldsymbol{x})d\boldsymbol{x}\, d\theta \\ &= \int_{\Theta} \{\int_{\chi} L(\theta, \delta(\boldsymbol{x}))g(\theta|\boldsymbol{x})d\theta\} f_{\boldsymbol{X}}(\boldsymbol{x})d\boldsymbol{x} \end{aligned}$$

انتگرال داخلی در رابطه اخیر، یعنی

$$\int_{\chi} L(\theta, \delta(\boldsymbol{x}))g(\theta|\boldsymbol{x})d\theta$$

مخاطره پسین نامیده می‌شود.

برآوردگر بیز پارامتر $\gamma(\theta)$ نسبت به چگالی پیشین g و تابع زیان L، که آن را برای سهولت با δ_g نمایش می‌دهیم، آن برآوردگر تعریف می‌شود که دارای مینیمم تابع مخاطره بیزی باشد،

یعنی

$$r(g,\delta_g) = \min r(g,\delta)$$

توجه داشته باشید که تعریف فوق روش پیدا کردن برآوردگر بیز $\gamma(\theta)$ را ارائه نمی‌کند. برای پیدا کردن برآوردگر بیز پارامتر $\gamma(\theta)$ ، می‌توانیم به صورت زیر عمل کنیم.
با توجه به این که به دنبال برآوردگری نظیر $\delta_g(\boldsymbol{X})$ هستیم که $r(g,\delta)$ را به عنوان تابعی از برآوردگرهای ممکن، یعنی به ازای هر $\delta \in D$، مینیمم کند و از طرفی نشان دادیم که:

$$\begin{aligned} r(g,\delta) &= \int_\Theta \{\int_\chi L(\theta,\delta(\boldsymbol{x}))f_\theta(\boldsymbol{x})d\boldsymbol{x}\}g(\theta)d\theta \\ &= \int_\Theta \{\int_\chi L(\theta,\delta(\boldsymbol{x}))g(\theta|\boldsymbol{x})d\theta\}f_{\boldsymbol{X}}(\boldsymbol{x})d\boldsymbol{x} \end{aligned}$$

و چون انتگرال دوگانه اخیر غیرمنفی است، انتگرال دوگانه را می‌توان مینیمم کرد، اگر بتوانیم عبارت داخل آکولاد، یعنی مخاطره پسین را برای هر $\boldsymbol{x}$ داده شده‌ای مینیمم کنیم، که آن را با $\delta_g(\boldsymbol{x})$ نمایش می‌دهیم. در صورتی که برای هر $\boldsymbol{x}$، $\delta_g(\boldsymbol{x})$ منحصربه‌فرد باشد، آن‌گاه برآوردگر بیز منحصربه‌فرد خواهد بود.
توجه داشته باشید که با تغییر توزیع پیشین و تابع زیان، برآوردگر بیز پارامتر مورد نظر تغییر می‌کند.

۶.۱ برآوردگرهای مینیماکس

همان طور که گفته شد، برآوردگر مینیماکس δ_m عبارت از برآوردگری است که سوپریمم مقدار تابع مخاطره آن کمتر یا مساوی سوپریمم مقدار تابع مخاطره هر برآوردگر دیگر باشد. یعنی δ_m یک برآوردگر مینیماکس است، اگر برای هر $\delta \in D$،

$$\sup_\theta R(\theta,\delta_m) \le \sup_\theta R(\theta,\delta)$$

استفاده از تعریف در دستیابی به برآوردگرهای مینیماکس چندان ساده به نظر نمی‌رسد. با استفاده از قضیه زیر، به عنوان ابزاری در دستیابی به برآوردگرهای مینیماکس می‌توان استفاده کرد.

قضیه ۱–۲: اگر $\delta_g(x)$ برآوردگر بیز پارامتر $\gamma(\theta)$ نسبت به چگالی پیشین $g(\theta)$ با تابع مخاطره ثابت باشد، آن‌گاه برآوردگر δ_g یک برآوردگر مینیماکس است.(اثبات در پیوست B)

فصل ۲

رگرسیون

۱.۲ مقدمه

تحلیل رگرسیونی یک روش آماری برای بررسی و به مدل درآوردن ارتباط بین متغیرهاست. کاربردهای رگرسیون متعدد بوده و معمولاً در زمینه‌های مختلف از جمله مهندسی، فیزیک، اقتصاد، مدیریت، علوم زیستی، پزشکی و علوم اجتماعی به کار می‌رود. در حقیقت تحلیل رگرسیونی شاید از جمله روش‌های آماری با بیشترین و وسیعترین کاربرد باشد.

۲.۲ رگرسیون و ساختن مدل

اولین بار سرفرانسیس گالتون [۱] در سال ۱۸۸۵ اصطلاح رگرسیون را در مقاله خود به کار برد و بعد از آن این موضوع در ابعاد وسیع‌تری بیان شد. تحلیل رگرسیونی در حقیقت یک ابزار آماری برای تشخیص رابطه بین یک یا چند متغیر مستقل با یک متغیر وابسته است.

[۱] Francis Galton

در رگرسیون درباره ارتباط بین متغیرها بحث می‌شود. متغیرها را می‌توان به دو دسته متغیرهای پیشگو[۲] یا مستقل و متغیرهای پاسخ [۳] یا وابسته تقسیم کرد. منظور از متغیرهای پیشگو یا مستقل متغیرهایی هستند که می‌توان آنها را برابر مقادیری گرفت که گرچه ممکن است تحت کنترل نباشند، اما می‌توانند مشاهده شوند و این متغیرها با توجه به نامشان به طور مستقل تغییر می‌کنند. در نتیجه تغییر عمدی این متغیرها اثرهایی در متغیرهای دیگر یعنی متغیرهای پاسخ پدید می‌آورند. وجه تمایز بین متغیرهای پیشگو و پاسخ همیشه روشن نیست و گاهی وابسته به هدف تحلیل‌گر است اما معمولاً در عمل می‌توان نقش متغیرها را تشخیص داد.

در رگرسیون معمولاً فرض بر این است که متغیرهای پیشگو مقید به تغییرات تصادفی نیستند ولی متغیر پاسخ تصادفی است. از نقطه‌نظر کاربردی این مطلب اغلب صحیح نیست اما به دلیل این که اگر متغیرهای پیشگو تصادفی باشند، شیوه برازش بسیار پیچیده می‌شود، برای اجتناب از این پیچیدگی معمولاً فرض می‌شود که تغییرات تصادفی هر یک از متغیرهای پیشگو در مقایسه با دامنه متغیر پیشگوی مشاهده شده آن قدر کوچک است که می‌توان از آن صرفنظر کرد.

باید توجه داشت که در رگرسیون رابطه بین متغیرهای پیشگو و پاسخ یک رابطه تابعی ریاضی نیست، بلکه رابطه‌ای آماری است. به این معنی که به ازای هر x مقدار y کاملاً مشخص و دقیق نیست و تعیین مقدار آن با مقداری خطا همراه است. نوع رابطه‌ای که ممکن است بین متغیرهای پیشگو و وابسته برقرار باشد، متنوع است اما معمولاً به دلیل سادگی و نیز به دلیل وجود برخی مبانی نظری در مورد تقریب روابط غیرخطی با روابط خطی، به یافتن رابطه خطی علاقه‌مندی بیشتری نشان داده می‌شود.

در حالت رگرسیون خطی مدل آماری به صورت $\boldsymbol{Y} = \boldsymbol{X}\,\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ درنظر گرفته می‌شود

---

[۲] Predictive

[۳] Response

که در آن $\boldsymbol{X}$ بردار متغیرهای پیشگو، β پارامترها و ε میزان خطا و $\boldsymbol{Y}$ متغیر پاسخ را نشان می‌دهد. پس از تشخیص و پذیرش نوع رابطه اولین مرحله برآورد پارامترهای نامعلوم مدل است. با برآورد کردن پارامترهای نامعلوم، مدل کاملاً مشخص شده و پس از انجام نیکویی برازش می‌توان از آن برای پیش‌بینی استفاده کرد.

۳.۲ کاربردهای مدل‌های رگرسیون

مدل‌های رگرسیونی برای مقاصدی چند مشتمل بر موارد زیر مورد استفاده قرار می‌گیرند.

۱— توصیف داده‌ها

۲— برآورد پارامترها

۳— پیشگویی و برآورد

۴— کنترل

۴.۲ رگرسیون خطی ساده

مدل رگرسیون خطی ساده مدلی است با یک متغیر رگرسیونی مستقل x که با یک متغیر پاسخ y ارتباطی به صورت خط مستقیم دارد. این مدل رگرسیونی خطی ساده عبارت است از:

$$y = \beta_0 + \beta_1 x + \varepsilon \qquad (۱.۲)$$

که β_0 عرض از مبداء و β_1 شیب خط ثابت‌های نامعلوم‌اند و ε مولفه خطای تصادفی است. فرض می‌شود خطاها دارای میانگین صفر و واریانس σ^2 نامعلوم هستند و بعلاوه مقادیر این خطا ناهمبسته‌اند یعنی این که مقدار یک خطا بستگی به مقدار هر خطای دیگر ندارد.

نکته قابل توجه این که مقادیر متغیر رگرسیونی x توسط محقق و تحلیل‌گر اختیار و با خطای قابل صرفنظر کردن اندازه‌گیری می‌شود. در حالی‌که به ازاء هر مقدار ممکن x، پاسخ y متغیری تصادفی است بدین معنی که به ازاء هر مقدار ممکن x متغیر تصادفی y دارای یک توزیع احتمال است. میانگین این توزیع عبارتست از:

$$E(y|x) = \beta_0 + \beta_1 x$$

و واریانس آن عبارتست از:

$$V(y|x) = V(\beta_0 + \beta_1 x + \varepsilon) = \sigma^2$$

بنابراین میانگین y تابعی خطی از x است در حالی‌که واریانس آن به مقدار x بستگی ندارد و بعلاوه چون خطاهای ε ناهمبسته‌اند پاسخ‌ها نیز ناهمبسته خواهند بود.
پارامترهای β_0 و β_1 ضرایب رگرسیون نامیده می‌شوند. β_1 (شیب) تغییر در میانگین توزیع y به ازاء یک واحد تغییر در x است. اگر دامنه x شامل $x = 0$ باشد، در این صورت β_0 (عرض از مبداء) میانگین پاسخ y برای $x = 0$ است. اگر دامنه x شامل صفر نباشد، در این صورت β_0 تعبیر عملی (واقعی) ندارد.

۵.۲ رگرسیون چندگانه خطی

مدل رگرسیونی که مشتمل بر بیش از یک متغیر رگرسیونی باشد مدل رگرسیون چندگانه نامیده می‌شود.
فرض می‌کنیم عمر مؤثر یک وسیله برش با سرعت و عمق بریدگی ارتباط دارد. یک مدل رگرسیون چندگانه که می‌تواند این ارتباط را توصیف کند عبارتست از:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \qquad (۲.۲)$$

که y نشان دهنده عمر مؤثر وسیله، x_1 نشان دهنده سرعت برش و x_2 عمق برش می‌باشد. این یک مدل رگرسیون خطی چندگانه با دو متغیر رگرسیونی x_1 و x_2 است. کاربرد اصطلاح خطی به لحاظ این است که رابطه (۲.۲) تابعی خطی از پارامترهای نامعلوم β_0، β_1 و β_2 می‌باشد. این مدل یک صفحه را در فضای متغیرهای رگرسیونی x_1، x_2 مشخص می‌کند که پارامتر β_0 ارتفاع عرض از مبداء صفحه رگرسیونی است. اگر دامنه تعریف داده‌ها شامل $x_1 = x_2 = 0$ باشد در این صورت β_0 میانگین y به شرط $x_1 = x_2 = 0$ خواهد بود. پارامتر β_1 حاکی از تغییر مورد انتظار برای y به ازاء یک واحد تغییر x_1 است، وقتی که x_2 ثابت است. همچنین β_2 برابر است با تغییر مورد انتظار برای y به ازاء یک واحد تغییر x_2 وقتی که x_1 ثابت نگهداشته شود.

در حالت کلی متغیر پاسخ y ممکن است به p متغیر رگرسیونی بستگی داشته باشد، مدل

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p + \varepsilon \qquad (۳.۲)$$

یک مدل رگرسیون چندگانه خطی با p متغیر رگرسیونی نامیده می‌شود. پارامترهای $\beta_j, j = 0, 1, \ldots, p$ ضرایب رگرسیون نامیده می‌شوند. این مدل یک ابرصفحه در فضای p بعدی از متغیرهای رگرسیونی x_j است. پارامتر β_j نشان دهنده تغییرات مورد انتظار متغیر پاسخ به ازاء یک واحد تغییر در x_j است، وقتی که همه متغیرهای رگرسیونی دیگر $x_j (i \neq j)$ ثابت نگهداشته شوند. به همین جهت پارامترهای $\beta_j, j = 0, 1, \ldots, p$ ضرایب جزئی رگرسیون نامیده می‌شوند.