بسم الله الرحمن الرحیم

دانشکده علوم

پایان نامه کارشناسی ارشد در رشته آمار ریاضی

عنوان پایان نامه

بررسی اندازه های هماهنگی در متغیرهای تصادفی دو بعدی گسسته

توسط:

نجمه احمدی

استاد راهنما :

دکتر زهره شیشه بر

۱۳۸۶/۱۱/۱

شهریور ماه ۱۳۸۶

به نام خدا

بررسی اندازه های هماهنگی در متغیرهای تصادفی دوبعدی گسسته

به وسیله‌ی :

نجمه احمدی

پایان‌نامه

ارائه‌شده به تحصیلات تکمیلی دانشگاه به‌عنوان بخشی

از فعالیت های لازم برای اخذ درجه کارشناسی ارشد

در رشته‌ی:

آمار ریاضی

دانشگاه شیراز

شیراز

جمهوری اسلامی ایران

ارزیابی شده توسط کمیته پایان‌نامه با درجه : عالی

دکتر زهره شیشه بر، استادیار بخش آمار (رئیس کمیته)..................................

دکتر عبدالرسول برهانی حقیقی،استادیار بخش آمار..................................

دکتر عبدالرضا بازرگان لاری، استادیار بخش آمار..................................

شهریورماه ۱۳۸۶

تقدیم به

پدر و مادر عزیزم

که تنها عظمت یاریشان توان پاهایم،انگیزه کارهایم و دلخوشی راهم بود و اینک ثمره دسترنج آنهاست که در وجودم متبلور است.

به **خواهران ، برادران و همسرم،**

و به تمام آنانی که

در مکتبشان علم آموختم.

به تمامی آنانی که

در مکتبشان علم آموختم.

سپاس مخصوص خدایی است که رحمت بی دریغش را زوالی نیست و دریای لطف و عنایتش را کرانه نا پیداست. هم او که هر چه هست از اوست و هر چه نیست مدیون اذن اوست.

اگر نتیجه این تلاش ، کورسویی از علم و دانش باشد و بتواند گامی هر چند نا محسوس در پیمودن مسیر علم به حساب آید، آن را مدیون بزرگوارانی می دانم که رسالت سنگین راهبری بر عهده آنان بوده هست . تشکر ویژه ام را به سرکار خانم دکتر شیشه بر تقدیم می کنم که در طول این دوره هم راهنمای زندگیم بوده اند ؛ هم در این تلاش مقدس ، چراغ راهم . همچنین از جناب آقای دکتر برهانی و جناب آقای دکتر بازرگان که در این راه مشاورم بودند، قدردانی می نمایم.

همچنین از همسر عزیزم که در این مدت از هیچ کمکی دریغ نورزید، او که برای هر اوج گرفتنی پر پروازم بود و برای هر سقوطی یک ناجی، قدردانی می کنم.

چکیده

بررسی اندازه‌های هماهنگی در متغیرهای تصادفی دو بعدی گسسته

به وسیله‌ی :

نجمه احمدی

روش ها و معیارهای زیادی برای توصیف و اندازه‌گیری ارتباط بین متغیرهای تصادفی وجود دارد. برای حالتی که متغیرها نرمال هستند معمولاً ضریب همبستگی گشتاوری پیرسن و ضریب های در ارتباط با آن به عنوان اندازه های ارتباطی مورد استفاده قرار می گیرند.

این پایان‌نامه به مطالعه ی اندازه‌های ارتباطی دیگر به ویژه آنهایی که به اندازه‌های هماهنگی ناپارامتری معروف هستند و شامل تو کندال، رو اسپیرمن،گامای جینی وفوترول اسپیرمن می باشند ،با استفاده از توابع مفصل خواهد پرداخت. برای این کار تابعی را معرفی می‌نماید که همیشه و به طور یکتا وجود دارد و به خوبی می‌تواند بیانگر کل ساختار وابستگی متغیرها باشد. این تابع که آن را تابع مفصل می‌نامند، در اینجا بررسی و خواص آن مرور خواهد شد. در ادامه اندازه‌های هماهنگی دو بعدی مورد بحث قرار می‌گیرند و ضمن معرفی کلاسی از توابع مفصل دو متغیره که از دو تابع یک متغیره ساخته می‌شود و دربرگیرنده خانواده‌های معروفی از توابع مفصل می‌باشد، اندازه‌های هماهنگی گوناگون را برای این کلاس محاسبه می‌نماید. همچنین ویژگیهای مختلف ارتباط، مفاهیم وابستگی و نظم جزئی و تعدادی مثال برای روشن تر شدن موضوع مورد بررسی قرار خواهد گرفت. به علاوه بررسی شرایط اندازه های وابستگی بر اساس هماهنگی در متغیرهای تصادفی دو بعدی گسسته انجام می گیرد که مهمترین تکنیک در این روش شامل توسیع پیوسته متغیرهای تصادفی گسسته بوسیله پیچش با متغیرتصادفی پیوسته ای می باشد که دارای تکیه گاه واحد است.

فهرست مطالب

عنوان	صفحه

عنوان	صفحه

فصل چهارم: بررسی انواع روابط وابستگی با استفاده از تابع مفصل

فصل پنجم: اندازه‌های هماهنگی در متغیرهای تصادفی دو بعدی گسسته

فهرست شکل ها

عنوان — **صفحه**

مقدمه

مقدمه

۱- معرفی

بررسی روابط وابستگی بین متغیرهای تصادفی یکی از موضوعاتی است که به طور گسترده در تئوری آمار و احتمال مورد مطالعه قرار گرفته است. روشها و معیارهای زیادی برای توصیف و اندازه‌گیری ارتباط بین متغیرها وجود دارد. برای حالتی که متغیرها نرمال هستند معمولاً ضریب همبستگی گشتاوری پیرسن و ضریبهای دیگر در ارتباط با آن بکار می‌روند. اما واضح است که فقط روابط وابستگی خطی می‌تواند توسط این روش به دست آید. اندازه‌های ارتباط[1] دیگری که به اندازه‌های هماهنگی ناپارامتری معروف هستند وجود دارد که میزان ارتباط بین متغیرها را مستقل از نوع توابع حاشیه‌ای نشان می‌دهند و با استفاده از تابع مفصل[2] به آسانی قابل بیان هستند و به خوبی می‌تواند بیانگر کل ساختار وابستگی متغیرها باشد. علاوه بر این ضرائب، انواع روابط نظم جزئی[3] و یکنواختی دم[4] نیز وجود دارند که به کمک آنها اطلاعات و جزئیات بیشتری از نوع ساختار تابع مفصل مربوط به متغیرها به دست می‌آید. به این ترتیب نتایج مفیدی از روابط موجود بین اندازه‌های ارتباط بیان می‌شود.

۲- مروری بر پژوهشهای پیشین

تابع مفصل در سال ۱۹۵۹ توسط اسکلار[5] معرفی گردید. او در قضیه معروف خویش که از مهمترین قضایا در مبحث توابع مفصل می‌باشد، وجود تابع مفصل را برای متغیرهای تصادفی اثبات کرد و نشان داد که این تابع برای متغیرهای تصادفی پیوسته یکتا خواهد بود.(برای مطالعه تاریخچه تابع مفصل به [۲۶] مراجعه کنید.)به دنبال معرفی تابع مفصل آماردانان به بررسی نقش آن در سایر مباحث آماری پرداختند که در فاصله زمانی ۱۹۵۹ تا ۱۹۷۶ بسیاری

1. Association measures
2. Copula
3. Partial ordering
4. Tail monotonicity
5. Sklar

از نتایج تحقیق مربوط به توسیع فضاهای متری برای توابع توزیع احتمال بود (اسکلار و شوازر[1] (۱۹۸۳)). رابطه وابستگی بین متغیرهای تصادفی با استفاده از ضرائب اندازه‌گیری ناپارامتری توسط شوازر و همکاران (۱۹۶۰) ؛ نلسن[2] و همکاران(۱۹۹۶) ؛ والف[3] و همکاران(۱۹۸۱)،فلورز[4] و همکاران(۲۰۰۱)، بررسی گردید و مشاهده شد که این ضرائب ناپارامتری به کمک توابع مفصل راحت تر قابل بیان هستند. امبلارد و گیرارد[5](۲۰۰۱)و لالنا[6] و فلورز(۲۰۰۴)خانواده‌های شبه‌پارامتری جدیدی از توابع مفصل دو متغیره را معرفی کردند و به بررسی ویژگیهای آنها پرداختند و نیز ضرائب ناپارامتری وابستگی را برایشان محاسبه کردند.

دنیوت و لامبرت[7](۲۰۰۵) کاربرد تابع مفصل برای متغیرهای تصادفی گسسته در زمینه اندازه‌های ارتباط از جمله τ کندال را مطالعه و نتایج مفیدی را بیان کردند.

۳- ساختار پایان‌نامه

در این پایان‌نامه، ارتباط بین تابع مفصل و اندازه‌های هماهنگی را بر اساس کار لالنا و فلورز[۲۲] ;دنیوت و لامبرت[۹] بررسی می‌کنیم و نشان خواهیم داد که اندازه‌های ارتباط، با استفاده از توابع مفصل به آسانی قابل بیان هستند.

در فصل اول (مفاهیم پایه) تعاریف و قضایای مورد نیاز در فصل‌های آتی را بیان خواهیم کرد. توجه اصلی ما در این فصل به معرفی تابع مفصل و قضیهٔ اسکلار که اساسی‌ترین قضیه این مبحث است و همچنین کلاس‌های معروفی از توابع مفصل می‌باشد.

در فصل دوم (معرفی ضرائب اندازه‌گیری غیر پارامتری) به معرفی و بررسی برخی خواص اندازه‌های ارتباط غیر پارامتری می‌پردازیم.

در فصل سوم (کلاس جدیدی از توابع مفصل) ضمن معرفی کلاس جدیدی از توابع مفصل به محاسبه ضرائب اندازه‌گیری ناپارامتری وابستگی در میان این کلاس خواهیم پرداخت.

---

1 Schwizer
2 Nelsen
3 Wolff
4 Flores
5 Ambelard & Girard
6 Lallena
7 Denuit & Lambert

در فصل چهارم (بررسی انواع وابستگی با استفاده از تابع مفصل) ویژگیهای نظم جزئی، انواع یکنوایی دم و وابستگی مربعی مثبت را در بین توابع مفصل و برای کلاس جدید بررسی خواهیم کرد.

در فصل پنجم (اندازه‌های هماهنگی در متغیرهای تصادفی دو بعدی گسسته)، به بررسی رفتار اندازه‌های وابستگی از جمله τ کندال برای داده‌های گسسته دو بعدی بر پایه هماهنگی خواهیم پرداخت.

فصل اول

مفاهیم پایه

۱- مفاهیم پایه

۱-۱- مقدمه

ارتباط بین توابع توزیع چند بعدی و توابع توزیع حاشیه‌ای با بعدهای کمتر آنها همواره مورد توجه آماردانان بوده است. فرچت[1] و دال اگلی[2] فعالیت‌های زیادی، در این زمینه انجام داده‌اند و ارتباط توابع توزیع دو متغیره و سه متغیره را با تابع توزیع حاشیه‌ای مطالعه کردند به دنبال این فعالیت‌ها، کلاس جدیدی از توابع، به نام تابع مفصل، برای بررسی ارتباط بین توابع توزیع حاشیه‌ای و تابع توزیع توأم یک متغیر تصادفی دوبعدی توسط اسکلار در سال ۱۹۵۹ معرفی شد. این کلاس جدید از توابع، در واقع توابع توزیع دو متغیره‌ای هستند که دارای تابع توزیع یکنواخت حاشیه‌ای می‌باشند.

در این فصل ابتدا تعاریف و قضایای مورد نیاز را ارائه خواهیم کرد و در ادامه به معرفی تابع مفصل و بیان و اثبات برخی از خواص آن می‌پردازیم. سپس چند نمونه از خانواده‌های معروف توابع مفصل را معرفی می‌نماییم.

۱-۲- تعاریف و مفاهیم پایه

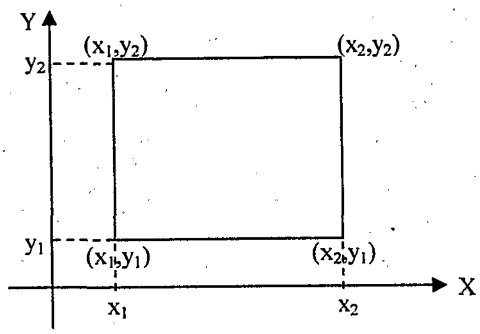
ابتدا برخی نمادگذاری‌ها و اصطلاحاتی که در این پایان‌نامه به کار خواهند رفت را ذکر می‌نماییم.

منظور از R، خط حقیقی معمولی $(-\infty, +\infty)$ و $\overline{R}$ خط حقیقی توسعه یافته $[-\infty,+\infty]$ و صفحه حقیقی توسعه یافته $\overline{R}\times\overline{R}$ می‌باشد. اگر B یک مستطیل در $\overline{R}^2$ باشد، به صورت زیر مشخص می‌شود:

$$B=[x_1,x_2]\times[y_1,y_2] \qquad (۱-۲-۱)$$

وقتیکه $x_1,x_2,y_1,y_2 \in R$ می باشند.

[1]. M. Frechet
[2]. G. Dall Agile



شکل ۱-۲-۱- مستطیل $B=[x_1,x_2]\times[y_1\times y_2]$

رئوس مستطیل B، نقاط (x_1,y_1)، (x_1,y_2)، (x_2,y_1)، (x_2,y_2) می‌باشند. مربع واحد I^2 نیز از ضرب خارجی $I\times I$ بوجود می‌آید که $I=[0,1]$ تابع حقیقی دو بعدی H، تابعی است که دامنه آن زیر مجموعه‌ای از $\overline{R}^2$ و برد آن زیر مجموعه R باشد.

توجه داشته باشید که در ادامه بحث گاهی اوقات دامنه تابع H را با نماد Dom H و برد آن را با نمادRan H نمایش داده‌ایم.

تعریف ۱-۲-۱- فرض کنید S_1 و S_2 زیر مجموعه‌های ناتهی از $\overline{R}$باشند و H تابعی باشد که دامنه آن $S_1\times S_2$ است، فرض کنید که $B=[x_1,x_2]\times[y_1,y_2]$ مستطیلی باشد که تمام رئوس آن در دامنه H قرار می‌گیرد، آنگاه "حجم B نسبت به H"[۱] را به صورت رابطه زیر تعریف می‌کنیم:

$$V_H(B)=H(x_2,y_2)-H(x_2,y_1)-H(x_1,y_2)+H(x_1,y_1) \qquad (۱-۲-۲)$$

تعریف ۱-۲-۲- یک تابع حقیقی دو بعدی H را " دو – صعودی "[۲] گوییم، هرگاه برای هر مستطیل B که رئوس آن متعلق به دامنه H است، $V_H(B)\geq 0$.

توجه: عبارت H تابعی دو – صعودی است به این معنا نخواهد بود که H در هر " شناسه "[۳] خود غیر نزولی است. عکس این موضوع نیز برقرار نمی‌باشد.

1 . H- Volume of B
2 . 2 - Increasing
3 . Argument

برای روشن شدن این مطلب به مثالهای زیر توجه کنید:

مثال ۱-۲-۱- فرض کنید H روی I^2 به صورت $H(x,y)=\max(x,y)$ تعریف شده باشد. بنابراین H تابعی غیر نزولی از x و y است، اما ۱- =$V_H\left(I^2\right)$ ، یعنی H، دو صعودی نیست.

مثال ۱-۲-۲- فرض کنید H روی I^2 به صورت زیر تعریف شده باشد:

$$H(x,y)=(2x-1)(2y-1)\ , \qquad x,y\in I$$

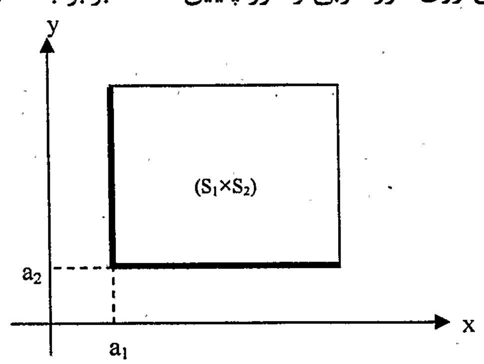
به آسانی می‌توان دید که H دو ــ صعودی است. اما H برای هر $y\in\left(0,\frac{1}{2}\right)$، یک تابع نزولی از x و همچنین برای هر $x\in\left(0,\frac{1}{2}\right)$ تابعی نزولی از y است.

لم ۱-۲-۱- فرض کنید S_1 و S_2 زیر مجموعه‌های ناتهی از $\overline{R}$ باشند و H یک تابع دو ــ صعودی با دامنه $S_1\times S_2$باشد. نقاط $x_1\le x_2$ را در S_1 و $y_1\le y_2$ را در S_2 در نظر بگیرید. نگاشت $t\mapsto H(t,y_2)-H(t,y_1)$ روی S_1 و نگاشت $t\mapsto H(x_2,t)-H(x_1,t)$ روی S_2 غیر نزولی می‌باشند.

به عنوان مثال کاربردی این لم ، نشان می‌دهیم که با یک فرض اضافی، یک تابع دو ــ صعودی روی هر مؤلفه‌ای غیر نزولی است. فرض کنید a_1 کوچکترین عضو S_1 و a_2 کوچکترین عضو S_2 باشد. می‌گوییم تابع H از $S_1\times S_2$ به روی R، "جهت‌دار"[1] است اگر برای هر $(x,y)\in S_1\times S_2$

$$H(a_1,y)=H(x,a_2)=0 \qquad (۳-۲-۱)$$

مثلاً اگر H تابعی جهت‌دار باشد و دامنه آن به صورتی باشد که در شکل (۱-۲-۲) نشان داده‌ایم، آنگاه مقدار تابع روی مرز غربی و مرز پایین دامنه برابر با صفر است.



شکل ۱-۲-۲- مثال مربوط به تعریف تابع جهت‌دار

[1] .grounded

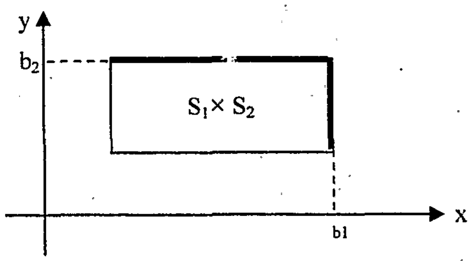
بنابراین خواهیم داشت:

لم ۱-۲-۲- فرض کنید $S_2 \times S_1$ زیر مجموعه‌های ناتهی از $\overline{R}$هستند اگر H تابعی دو - صعودی و جهت‌دار با دامنه $S_1 \times S_2$ باشد، آنگاه H در هر شناسه اش غیر نزولی است.

اثبات: فرض کنید، a_1 کوچکترین عضو S_1 و a_2 کوچکترین عضو S_2 باشد، در لم (۱-۲-۱) قرار دهید $x_1=a_1$ و $x_2=a_2$، اکنون فرض کنید b_1 بزرگترین عضو S_1 و b_2 بزرگترین عضو S_2 باشد، آنگاه می‌گوییم تابع H از $S_1 \times S_2$ به روی R، دارای توابع حاشیه‌ای F و G است اگر:

$$DomF = S_1 \quad , \quad F(x) = H(x, b_2) \quad \forall x \in S_1$$
$$DomG = S_2 \quad , \quad G(y) = H(b_1, y) \quad \forall y \in S_2$$

به عنوان مثال فرض کنید دامنه H به صورت ناحیه مشخص شده در شکل (۱-۲-۳) باشد، آنگاه F برابر است با مقدار تابع H روی مرز بالایی دامنه و G برابر است با مقدار تابع H روی مرز شرقی دامنه.



شکل ۱-۲-۳- مثال مربوط به تعریف تابع حاشیه‌ای

لم ۱-۲-۳- فرض کنید S_1 و S_2 زیر مجموعه‌های ناتهی از $\overline{R}$ هستند و H تابعی دو – صعودی جهت‌دار است با توابع حاشیه‌ای که دامنه آن $S_1 \times S_2$ قرار دارد. اگر (x_1, y_1) و (x_2, y_2) دو نقطه دلخواه در $S_1 \times S_2$ باشند آنگاه:

$$|H(x_2, y_2) - H(x_1, y_1)| \le |F(x_2) - F(x_1)| + |G(y_2) - G(y_1)| \qquad (۱-۲-۴)$$

اثبات: بنابر نامساوی مثلثی داریم:

$$|H(x_2, y_2) - H(x_1, y_1)| \le |H(x_2, y_2) - H(x_1, y_2)| + |H(x_1, y_2) - H(x_1, y_1)|$$

حال فرض کنید $x_1 \le x_2$ چون H جهت‌دار، دو صعودی و دارای توابع حاشیه‌ای است، با استفاده از لم‌های (۱-۲-۱) و (۱-۲-۲) داریم که:

$$0 \le H(x_2, y_2) - H(x_1, y_2) - H(x_1, y_2) \le F(x_2) - F(x_1)$$

وقتی $x_2 \le x_1$ نامساوی مشابهی برقرار است. بنابراین برای هر x_1 و x_2 در S_1 داریم:

$$|H(x_2, y_2) - H(x_1, y_2)| \le |F(x_2) - F(x_1)|$$

به طور مشابه برای هر y_1 و y_2 در S_2 داریم:

$$|H(x_1, y_2) - H(x_1, y_1)| \le |G(y_2) - G(y_1)|$$

پس:

$$|H(x_2, y_2) - H(x_1, y_1)| \le |F(x_2) - F(x_1)| + |G(y_2) - G(y_1)|$$

تعریف ۱-۲-۴- فرض کنید F یک تابع توزیع باشد، آنگاه شبه معکوس F، با دامنه I تابع $F^{(-1)}$ خواهد بود بطوریکه:

۱. اگر $t \in RanF$، آنگاه $F^{(-1)}(t)$ عدد $x \in \overline{R}$ است جائیکه $F(x) = t$ یعنی برای هر t در برد F، $F(F^{(-1)}(t)) = t$.

۲. اگر $t \notin RanF$، آنگاه $F^{(-1)}(t) = \inf\{x | F(x) \ge t|\} = Sup\{x \mid F(x) \le t\}$

اگر F یک تابع اکیداً صعودی باشد، آنگاه یک شبه معکوس دارد که همان تابع معکوس معمولی است که معمولاً آن را با F^{-1} نمایش می‌دهیم.

مثال ۱-۲-۳- اگر برای هر a در R، تابع پله‌ای $\varepsilon_a(.)$ را به صورت:

$$\varepsilon_a(x) = \begin{cases} 1, & x \in [-\infty, a) \\ 0, & x \in [a, +\infty] \end{cases}$$

تعریف کنیم آنگاه شبه معکوس ε_a به صورت:

$$\varepsilon_a^{(-1)}(t) = \begin{cases} a_0 & t = 0 \\ a & t \in (0,1) \\ a_1 & t = 1 \end{cases}$$

که a_0 و a_1 دو عدد در $\overline{R}$ هستند بطوریکه $a_0 \le a \le a_1$.
از مثال فوق نتیجه می‌شود که ε_a^{-1} یکتا نیست.

تعریف ۱-۲-۵- فرض کنید φ یک تابع پیوسته اکیداً نزولی با دامنه I و برد $[0,\infty]$ باشد بطوریکه $\varphi(1)=0$. آنگاه معکوس — مانند φ تابع $\varphi^{[-1]}$ خواهد بود با $Dom\,\varphi^{[-1]}$ و $Ran\varphi = I$ بطوریکه

$$\varphi^{[-1]}_{(t)} = \begin{cases} \varphi^{-1}(t) & 0 \le t \le \varphi(0) \\ 0 & \varphi(0) \le t \le \infty \end{cases}$$

توجه داشته باشید که $\varphi^{[-1]}$ روی $[0,\infty]$ پیوسته و غیر صعودی و روی $[0,\varphi(0)]$ اکیداً نزولی است همچنین روی I، $\varphi^{[-1]}(\varphi(u)) = u$

$$\varphi\left(\varphi^{-1}(t)\right) = \begin{cases} t & 0 \le t \le \varphi(0) \\ \varphi(0) & \varphi(0) \le t \le \infty \end{cases} = \min\left(t, \varphi(0)\right)$$

و بالاخره، اگر $\varphi(0)=\infty$ آنگاه $\varphi^{[-1]} = \varphi^{-1}$.

تعریف ۱-۲-۶- فرض کنید بازه $I \subset R$ و تابع $f : I \to R$ تعریف شده باشد. f یک تابع محدب روی I است اگر

$$f((1-t)a + tb) \le (1-t)f(a) + tf(b) \qquad (۵-۲-۱)$$

بطوریکه $a,b \in I$ و $0 \le t \le 1$ و تابع f را نیز مقعر گوییم اگر f تابعی محدب باشد.

تعریف ۱-۲-۷- فرض کنید $[a,b] \subseteq R$ باشد تابع $f : [a,b] \to R$ پیوسته مطلق روی $[a,b]$ است، اگر برای هر $\varepsilon>0$، وجود داشته باشد $\delta > 0$ بطوریکه اگر $\{(a_j,b_j)\}_{j=1}^{n}$ خانواده‌ای از زیر بازه‌های جدا از هم درون [a,b] باشد، آنگاه $\sum_{j=1}^{n}(b_j - a_j) < \delta$ نتیجه دهد $\sum_{j=1}^{n}\left|f(b_j) - f(a_j)\right| < \varepsilon$.

قضیهٔ ۱-۲-۱- فرض کنید $f : [a,b] \to R, [a,b] \subset R$ ، تابعی پیوسته و محدب یا مقعر روی بازه [a,b] باشد. آنگاه f نیز یک تابع پیوسته مطلق روی بازه [a,b] می‌باشد. [28].