

(۱) زنجیره مارکوف (Markov chain)

تعریف ۱- بردار $u = (u_1, u_2, \dots, u_n) \neq 0$ بردار ثابت (fixed vector) یا بردار ثابتی ماتریس مربع $A_{n \times n}$ است، هرگاه:

$$uA = u \quad (۱)$$

واضح است که اگر u بردار ثابت ماتریس A باشد، آنگاه به ازای هر اسکالر k ($k \neq 0$)، بردار ku هم چنین است.

مثال: $u = (1, -2)$ ، بردار ثابت ماتریس $A = \begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix}$ است.

تعریف ۲- بردار $u = (u_1, u_2, \dots, u_n)$ را یک بردار احتمال (probability vector) نامیم، هرگاه مؤلفه‌های آن غیر منفی و مجموع آنها واحد باشد.

تعریف ۳- ماتریس مربع $P = (p_{ij})$ را یک ماتریس استوکاستیک (ماتریس تصادفی) (stochastic matrix) نامیم، هرگاه هر سطر آن یک بردار احتمال باشد.

قضیه ۱- اگر A و B دو ماتریس (مربع) تصادفی باشند، حاصلضرب آنها (AB) نیز یک ماتریس تصادفی است. در حالت خاص A^n ($n \in \mathbb{N}$) هم یک ماتریس تصادفی است.

اثبات: فرض کنید $A \cdot B = C$ ، داریم $C_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$ ، لذا $C_{ij} \geq 0$ و هم چنین مجموع درایه‌های سطر نام ماتریس C :

$$\sum_{j=1}^n C_{ij} = \sum_{j=1}^n \sum_{k=1}^n a_{ik} b_{kj} = \sum_{k=1}^n \sum_{j=1}^n a_{ik} b_{kj} = \sum_{k=1}^n a_{ik} \left(\sum_{j=1}^n b_{kj} \right)$$

ولی: $\sum_{j=1}^n b_{kj} = 1$ و $\sum_{k=1}^n a_{ik} = 1$ ، لذا:

$$\sum_{j=1}^n C_{ij} = \sum_{k=1}^n a_{ik} (1) = \sum_{k=1}^n a_{ik} = 1$$

□

(۲)
تعریف ۱- ماتریس تصادفی P را منظم (regular) نامیده می‌گویند اگر
درایه‌های آن توانی از آن (مثلاً P^m و $m \in \mathbb{N}$) مثبت باشند

مثال ۱: ماتریس تصادفی $A = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ منظم نیست زیرا:

$$A^2 = \begin{pmatrix} 1 & 0 \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}, A^3 = \begin{pmatrix} 1 & 0 \\ \frac{7}{8} & \frac{1}{8} \end{pmatrix}, A^4 = \begin{pmatrix} 1 & 0 \\ \frac{15}{16} & \frac{1}{16} \end{pmatrix}, \dots, A^k = \begin{pmatrix} 1 & 0 \\ \frac{2^k-1}{2^k} & \frac{1}{2^k} \end{pmatrix}$$

در صورتی که ماتریس $B = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ منظم است، زیرا:

$$B^2 = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

قضیه ۲- اگر P یک ماتریس تصادفی و u یک بردار احتمال باشد، آنگاه
 uP یک بردار احتمال است.

اثبات. با فرض $v = uP$ ، داریم $v_j = \sum_{i=1}^n u_i p_{ij}$ ، داریم $v_j \geq 0$ و

$$\sum_{j=1}^n v_j = \sum_{j=1}^n \sum_{i=1}^n u_i p_{ij} = \sum_{i=1}^n \sum_{j=1}^n u_i p_{ij} = \sum_{i=1}^n u_i \left(\sum_{j=1}^n p_{ij} \right)$$

$$\sum_{j=1}^n p_{ij} = 1 \quad \text{ولی} \quad \text{لذا:} \quad \sum_{j=1}^n v_j = 1$$

قضیه زیر همین خاصیت اساسی ماتریس‌های تصادفی منظم است (بدون اثبات)

قضیه ۳- اگر P یک ماتریس تصادفی منظم باشد، آنگاه:

۱- P دارای یک بردار ثابت احتمال منحصر به فرد t است که نام توانی آن مثبت هستند.

۲- دنباله توانهای P یعنی $P, P^2, P^3, P^4, \dots$ به سمت ماتریس T

میل می‌کند که هر سطر آن همان ماتریس ثابت t است.

۳- اگر p یک بردار احتمال باشد، آنگاه دنباله بردارهای

$p, pP, pP^2, pP^3, \dots$ به سمت بردار ثابت t میل می‌کند.

(و حد این دنباله مستقل از p است)

(۳)

توجه: وقتی می‌گوئیم P^n به پست T می‌کند، این یعنی راهی دهد که بر درایه P^n به پست درایه تناظرش از T می‌کند، همین توصیف هم برای P صورت می‌گیرد.
سوال ۲- ماتریس تصادفی:

$$P = \begin{pmatrix} 0 & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{ (از نظری گرام)}$$

دید می‌شود که:

$$P^2 = \begin{pmatrix} \frac{3}{8} & \frac{5}{8} & 0 \\ \frac{1}{4} & \frac{5}{8} & \frac{1}{8} \\ \frac{1}{4} & \frac{1}{4} & 0 \end{pmatrix}, \quad P^3 = \begin{pmatrix} \frac{5}{16} & \frac{19}{32} & \frac{3}{32} \\ \frac{5}{16} & \frac{5}{8} & \frac{1}{16} \\ \frac{1}{4} & \frac{5}{8} & \frac{1}{8} \end{pmatrix}$$

تمام درایه‌های P^3 مثبت هستند، پس P یک ماتریس تصادفی منظم است و دارای یک بردار احتمال مثبت منحصر به فرد است که اگر آن را $u = (x, y, 1-x-y)$ بگیریم، باید داشته باشیم:

$$(x, y, 1-x-y) \begin{pmatrix} 0 & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 1 & 0 \end{pmatrix} = (x, y, 1-x-y)$$

$$\Rightarrow \begin{cases} \frac{1}{4}y = x \\ \frac{3}{4}x + \frac{1}{4}y + 1 - x - y = y \\ \frac{1}{4}x = 1 - x - y \end{cases} \Rightarrow \begin{cases} x = \frac{4}{13} \\ y = \frac{1}{13} \end{cases} \Rightarrow u = \left(\frac{4}{13}, \frac{1}{13}, \frac{1}{13} \right)$$

از طرف دیگر:

$$P^4 = \begin{pmatrix} \frac{19}{64} & \frac{5}{8} & \frac{5}{64} \\ \frac{5}{16} & \frac{39}{64} & \frac{5}{64} \\ \frac{5}{16} & \frac{5}{8} & \frac{1}{16} \end{pmatrix} = \begin{pmatrix} 0.2969 & 0.6250 & 0.0781 \\ 0.3125 & 0.6094 & 0.0781 \\ 0.3125 & 0.6250 & 0.0625 \end{pmatrix}$$

$$P^5 = \begin{pmatrix} \frac{5}{19} & \frac{157}{259} & \frac{19}{259} \\ \frac{39}{128} & \frac{79}{128} & \frac{5}{64} \\ \frac{5}{19} & \frac{39}{64} & \frac{5}{64} \end{pmatrix} = \begin{pmatrix} 0.2625 & 0.6062 & 0.0742 \\ 0.3047 & 0.6172 & 0.0781 \\ 0.2625 & 0.6094 & 0.0781 \end{pmatrix} \quad (4)$$

$$P^n \rightarrow T = \begin{pmatrix} \frac{4}{13} & \frac{1}{13} & \frac{1}{13} \\ \frac{4}{13} & \frac{1}{13} & \frac{1}{13} \\ \frac{4}{13} & \frac{1}{13} & \frac{1}{13} \end{pmatrix} = \begin{pmatrix} 0.3077 & 0.0769 & 0.0769 \\ 0.3077 & 0.0769 & 0.0769 \\ 0.3077 & 0.0769 & 0.0769 \end{pmatrix}$$

برای بردار احتمال P خطی به نام T داریم که از T است

معنی:

$$P P^n \rightarrow \mu : p = (p_1 \ p_2 \ p_3)$$

$$(p_1 \ p_2 \ p_3) \begin{pmatrix} \frac{4}{13} & \frac{1}{13} & \frac{1}{13} \\ \frac{4}{13} & \frac{1}{13} & \frac{1}{13} \\ \frac{4}{13} & \frac{1}{13} & \frac{1}{13} \end{pmatrix} = \left(\frac{4}{13} \ \frac{1}{13} \ \frac{1}{13} \right)$$

(زیرا: $p_1 + p_2 + p_3 = 1$)

زنجیره‌های مارکوف (Markov Chains)

تعریف ۵- یک دنباله آزمایشها (trials) یا نتایج یا خروجی (outcomes) $X_1, X_2, \dots$ که در دو شرط زیر صدق می‌کند در نظر می‌گیریم:

۱- نتیجه هر آزمایش متعلق به مجموعه‌تنتاهی $\{a_1, a_2, \dots, a_m\}$ است که آن را فضای حالت (state space) دستگاه می‌نامیم. اگر نتیجه آزمایش n -ام دستگاه a_i باشد، می‌گوییم دستگاه در زمان n یا در گام n -ام در حالت a_i است.

۲- نتیجه هر آزمایش حد اکثر وابسته به نتیجه آزمایش یک گام قبل از آن است و مستقل از نتایج آزمایش‌های قبلی دیگر است.

(۵)

به ازای هر زوج از حالت i کی (a_i, q_i) ، احتمال p_{ij} تناظر است و آن احتمال این است که حالت j بلافاصله بعد از حالت i به وقوع پیوندد .
چنین فرایند (process) تصادفی را پیکره مارکوف متناهی (finite) می نامند . اعداد p_{ij} را احتمالات انتقال (transition probabilities) نامیده می شود که می توان در ماتریس زیر آرایش داد

$$P = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1m} \\ p_{21} & p_{22} & \dots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \dots & p_{mm} \end{pmatrix}$$

که آن را ماتریس انتقال یا برین با هر حالت i ، سطر i -ام $(p_{i1}, p_{i2}, p_{i3}, \dots, p_{im})$ می نامند .

تناظر است ، بطوریکه اگر سیستم در حالت i باشد ، آنگاه این سطر احتمالات وقوع تمام حالت های ممکن بلافاصله بعد از i را نشان می دهد لذا این یک بردار احتمال است . و از آنجا :

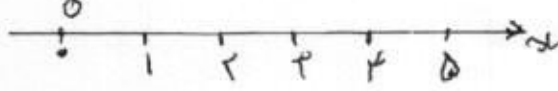
قضیه ۴ - ماتریس انتقال P در یک پیکره مارکوف یک ماتریس استوکیاست .

مثال ۳ - شخصی هر روز یا با قطار و یا با ماشین (شخصی) به سرکاری می رود . بطوریکه هیچ دو روز متوالی از قطار استفاده نمی کند و اگر روزی با ماشین سرکار رود فردای آن روز با احتمالات یکسان از قطار و ماشین استفاده می کند . چنین فرآیندی یک پیکره مارکوف است ، فضای حالت $\{(a, d), (a, t)\}$ با ماتریس انتقال

$$P = \begin{pmatrix} t & d \\ 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \text{ است .}$$

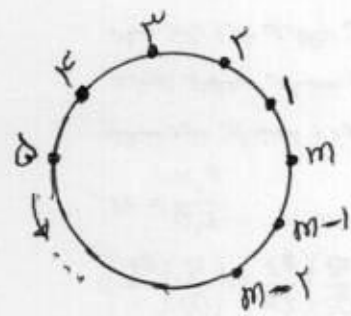
مثال ۴ - ترمک A به تصادف در نقاط به طول های ۵ و ۲ و ۱ و ۰ روی

محور xx ، در فواصل زمانی گام بر میدارد. اگر در نقطه ای در یکی از نقاط a_0 باشد با احتمال p به نقطه مجاور در جهت راست و با احتمال q به نقطه مجاور در جهت چپ میرود بطوریکه $p+q=1$. در صورتی که در نقطه 0 باشد به نقطه 1 و اگر در نقطه 5 باشد به نقطه 4 گام میزند. چنین فرایندی یک زنجیره مارکوف است که در آن فضای حالت: $\{a_0, a_1, \dots, a_4, a_5\}$ است، که a_0 بین بود و a_5 حرکت در نقطه 5 است. ماتریس انتقالی این فرایند عبارت است از:

$$P = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & a_5 \\ a_0 & q & 0 & p & 0 & 0 \\ a_1 & p & 0 & q & 0 & 0 \\ a_2 & 0 & q & 0 & p & 0 \\ a_3 & 0 & 0 & q & 0 & p \\ a_4 & 0 & 0 & 0 & q & 0 \\ a_5 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$


مثال ۵ -

m نقطه (جایگاه) $1, 2, 3, \dots, m$ روی محیط دایره ای که در جهت خلاف حرکت عقربه های ساعت بر چرخش گذاری شده است، مفروض است. یک شکر که این نقاط را طی می کند، در هر نقطه که باشد با احتمال p خلاف جهت حرکت عقربه ساعت و با احتمال q در جهت حرکت عقربه ساعت به نقطه مجاور گام بر میدارد. این فرایند یک زنجیره مارکوف با فضای حالت $\{1, 2, 3, \dots, m\}$ است و ماتریس انتقالی این فرایند:



$$P = \begin{pmatrix} 1 & 2 & 3 & \dots & m-2 & m-1 & m \\ 1 & q & 0 & \dots & 0 & 0 & p \\ 2 & p & q & \dots & 0 & 0 & 0 \\ 3 & 0 & p & q & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ m-1 & 0 & 0 & \dots & q & 0 & p \\ m & p & 0 & \dots & 0 & q & 0 \end{pmatrix}$$

احتمالات انتقال مراتب بالاتر

درایه p_{ij} در ماکریس انتقال P عبارتست از احتمال تغییر سیستم از حالت i به حالت j در یک گام. سؤال: احتمال اینکه سیستم از حالت i پس از n گام به حالت j تغییر وضعیت دهد، یعنی دانسته باشیم:

$$a_i \rightarrow a_{k_1} \rightarrow a_{k_2} \rightarrow \dots \rightarrow a_{k_{n-1}} \rightarrow a_j$$

تعدد است؟

اگر این احتمال را به $p_{ij}^{(n)}$ نشان دهیم و این درایه یک ماکریس $p^{(n)}$ باشد، ماکریس $p^{(n)}$ را ماکریس انتقال n گام (n -step transition matrix) می‌نامیم. قضیه ۶- اگر P ماکریس انتقال یک زنجیره مارکوف باشد، در این صورت ماکریس انتقال n گام، برابر است با $p^{(n)} = P^n$.

فرض کنید در یک زمان دلخواه احتمال اینکه سیستم در وضعیت i باشد برابر p_i باشد ($i=1, 2, \dots, m$) در این صورت بردار $p = (p_1, p_2, \dots, p_m)$ را توزیع احتمال (probability distribution) سیستم در آن لحظه‌مانند در حالت خاص

سیستم و $p^{(0)} = (p_1^{(0)}, p_2^{(0)}, \dots, p_m^{(0)})$ ، نالین توزیع احتمال اولیه n -گام نالین می‌دهیم. $p^{(n)} = (p_1^{(n)}, p_2^{(n)}, \dots, p_m^{(n)})$ را توزیع احتمال سیستم در گام

قضیه ۷- اگر P ماکریس انتقال یک زنجیره مارکوف و $p = (p_1, \dots, p_m)$ توزیع احتمال سیستم در لحظه‌ای دلخواه باشد، در این صورت $p P^n$ توزیع احتمال سیستم در گام بعدی و $p P^n$ توزیع احتمال سیستم در n گام بعدی است در حالت خاص

$$p^{(1)} = p^{(0)} P, p^{(2)} = p^{(1)} P, p^{(3)} = p^{(2)} P, \dots, p^{(n)} = p^{(n-1)} P = p^{(0)} P^n$$

مثال ۶- زنجیره مارکوف مثال (۱) را در نظر می‌گیریم که ماکریس انتقال آن می‌باشد، که t بیاگر رفتن یا برگشتن و d بیاگر رفتن به

$$P = \begin{pmatrix} t & d \\ \frac{1}{t} & \frac{1}{d} \end{pmatrix}$$

به سرکار با ماشین است. حال: (۸)

$$P^4 = P^2 P^2 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix} = \begin{pmatrix} \frac{3}{8} & \frac{5}{8} \\ \frac{5}{16} & \frac{11}{16} \end{pmatrix}$$

پس احتمال اینکه سیستم از حالت t پس از n گام به حالت d تغییر کند $P_{td}^{(n)}$ است، مثلاً

$$P_{tt}^{(4)} = \frac{3}{8}, \quad P_{dt}^{(4)} = \frac{5}{16}, \quad P_{dd}^{(4)} = \frac{11}{16} \quad \text{(قضیه ۶)}$$

حال اگر احتمال اینکه شخصی در روز اول با ترک سرکار برود $\frac{5}{9}$ و با بماندن $\frac{4}{9}$ باشد، توزیع احتمال دو وضعیت (t, d) در روز پنجم عبارتست از

$$P^{(5)} = P^{(4)} P = \left(\frac{5}{9}, \frac{4}{9} \right) \begin{pmatrix} \frac{3}{8} & \frac{5}{8} \\ \frac{5}{16} & \frac{11}{16} \end{pmatrix} = \left(\frac{35}{96}, \frac{41}{96} \right)$$

یعنی با احتمال $\frac{35}{96}$ روز پنجم با ترک و با احتمال $\frac{41}{96}$ با بماندن به سرکار برود (قضیه ۷)

توزیع مانا در یک زنجیره مارکوف منظم

stationary distribution of regular markov chain

فرض کنید یک زنجیره مارکوف منظم است، یعنی ماتریس انتقال P آن یک ماتریس منظم است، در این صورت ماتریس انتقال n -گام یعنی P^n به سمت ماتریس π میل می کند که تمام سطرهاي آن بردار t یعنی بردار یعنی بردار احتمال پیشین (یابست) ماتریس P می باشد. حال $P_{ij}^{(n)}$ که عبارتست از احتمال تغییر وضعیت سیستم از حالت i به حالت j در تعداد n گام های به اندازه کافی بزرگ n مستقل از حالت اولیه i است، و به سمت شولف π_j از بردار t میل می کند

قضیه ۸ - اگر ماتریس انتقال P یک زنجیره مارکوف منظم باشد، آنگاه در طی آزمایشات به تعداد دفعات زیاد، احتمال وقوع حالت j برابر است با π_j یعنی مؤلفه j ام ماتریس ثابت π ماتریس P .

ماتریس π را توزیع مانا (stationary distribution) در زنجیره مارکوف می نامیم.

مثال ۷ - در مثال ۶ (ترک - ماشین) $P = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ و به فرض

$t = (x, 1-x)$ ، آنگاه $(x, 1-x) \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = (x, 1-x)$ و از اینجا:

$x = \frac{1}{3}$ و $t = (\frac{1}{3}, \frac{2}{3})$ یعنی در طولانی مدت، شخصی ۳۳٪ اوقات با ترک و ۶۷٪ اوقات با ماشین به سر کار می رود.

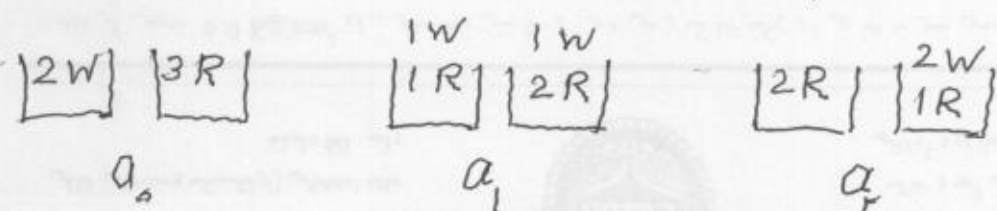
مثال ۸ - ۲ دوره سفید درجه A و ۳ دوره قرمز درجه B قرار دارد. دوره ۲ یک اندازه هستند. برابر به تصادف یک دوره از هر درجه را انتخاب کرده (بدون توجه به رنگ آنها) و جای دوره انتخاب شده را عوض می کنیم. اگر حالت a_i نشان دهند، ۲ دوره قرمز درجه A باشد:

۱ - ماتریس انتقال این فرایند را بدست آورید.

۲ - احتمال اینکه بعد از سه گام، ۲ دوره قرمز درجه A باشد چقدر است؟

۳ - در آزمایش های طولانی مدت، احتمال اینکه ۲ دوره قرمز در A باشد چقدر است؟

حل - فضای حالت: $\{a_0, a_1, a_2\}$



(۱۰)

$$p_{0,0} = 0$$

$$p_{0,1} = P(a_0 \rightarrow a_1) = 1$$

$$p_{0,2} = P(a_0 \rightarrow a_2) = 0$$

$$p_{1,0} = P(a_1 \rightarrow a_0) = \frac{1}{3} \times \frac{1}{3} = \frac{1}{9}$$

$$p_{1,1} = P(a_1 \rightarrow a_1) = \frac{1}{3} \times \frac{1}{3} + \frac{1}{3} \times \frac{2}{3} = \frac{1}{3}$$

$$p_{1,2} = P(a_1 \rightarrow a_2) = \frac{2}{3} \times \frac{1}{3} = \frac{1}{3}$$

$$p_{2,0} = P(a_2 \rightarrow a_0) = 0$$

$$p_{2,1} = P(a_2 \rightarrow a_1) = \frac{2}{3}$$

$$p_{2,2} = P(a_2 \rightarrow a_2) = \frac{1}{3}$$

$$P = \begin{matrix} & a_0 & a_1 & a_2 \\ \begin{matrix} a_0 \\ a_1 \\ a_2 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{9} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{pmatrix} \end{matrix}$$

ماتریس انتقال :

فداریم :

$$P^2 = \begin{pmatrix} \frac{1}{9} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{12} & \frac{23}{36} & \frac{5}{18} \\ \frac{1}{9} & \frac{5}{9} & \frac{1}{3} \end{pmatrix}$$

در عناصر P^2 ثبت می‌شود
لذا P منتظم است

۲- چون سیستم از حالت a_0 شروع می‌شود لذا :

$$p^{(0)} = (1, 0, 0)$$

$$\Rightarrow p^{(1)} = (1, 0, 0) P = (0, 1, 0), \quad p^{(2)} = p^{(1)} P = (0, 1, 0) P = \left(\frac{1}{9}, \frac{1}{3}, \frac{1}{3}\right)$$

$$p^{(3)} = p^{(2)} P = \left(\frac{1}{9}, \frac{1}{3}, \frac{1}{3}\right) P = \left(\frac{1}{12}, \frac{23}{36}, \frac{5}{18}\right) \Rightarrow p_2^{(3)} = \frac{5}{18}$$

جواب

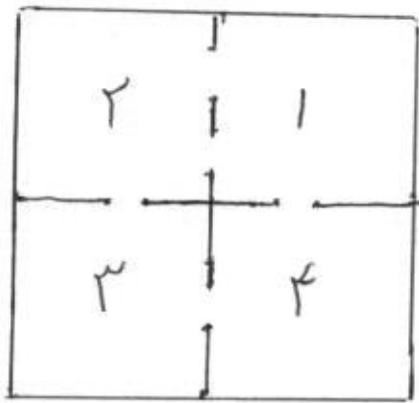
۳- با فرض $t = (x, y, 1-x-y)$ بردار ثابتی ماتریس P

$$\Rightarrow (x, y, 1-x-y) \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{9} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{pmatrix} = (x, y, 1-x-y) \Rightarrow \begin{cases} \frac{1}{9}y = x \\ x + \frac{1}{9}y + \frac{2}{3}(1-x-y) = y \\ \frac{1}{9}y + \frac{1}{3}(1-x-y) = 1-x-y \end{cases}$$

$$\Rightarrow x = \frac{1}{10}, y = \frac{9}{10} \Rightarrow t = \left(\frac{1}{10}, \frac{9}{10}, \frac{3}{10}\right)$$

پس در طولانی مدت $\frac{1}{10}$ اوقات ۲ عدد در قرمز درج می‌گردد
قراری می‌گردد

شکل زیر و ماتریس M اطاق که بوسید در هائیک
به هم مرتبط اند، می باشد.



موشی با عبور از درب ها، با احتمالات مساوی
بطور تصادفی از یک اطاق به اطاق دیگر می رود
در یک زمان طولانی احتمال بودن موش در هر

اطاق چقدر است. حل - فضای حالت: $\{a_1, a_2, a_3, a_4\}$

a_i ماتریس بودن موش در اطاق شماره i . مارکس انتقال:

$$P = \begin{matrix} & a_1 & a_2 & a_3 & a_4 \\ \begin{matrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{matrix} & \begin{pmatrix} 0 & \frac{2}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \end{matrix} \Rightarrow P^2 = \begin{pmatrix} \frac{4}{9} & 0 & \frac{7}{18} & \frac{1}{9} \\ 0 & \frac{11}{18} & 0 & \frac{7}{18} \\ \frac{1}{3} & 0 & \frac{5}{12} & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{3} \end{pmatrix}$$

دیده می شود که تمام عناصر مارکس $P^4 = P^2 \times P^2$ همگی مثبت هستند، لذا مارکس
استوکاستیک P منظم است و دارای یک بردار ثابتی احتمال:

$t = (x, y, z, 1-x-y-z)$ با $1-x-y-z > 0$ است.

$$(x, y, z, 1-x-y-z) \begin{pmatrix} 0 & \frac{2}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} \end{pmatrix} = (x, y, z, 1-x-y-z)$$

$$\begin{cases} \frac{2}{3}y = x \\ \frac{1}{3}x + \frac{1}{3}z = y \\ \frac{1}{3}y + \frac{1}{3}(1-x-y-z) = z \end{cases} \Rightarrow \begin{cases} x = \frac{2}{13} \\ y = \frac{3}{13} \\ z = \frac{10}{39} \end{cases}$$

$$\Rightarrow x = \frac{2}{13}, y = \frac{3}{13}, z = \frac{10}{39}$$

$$t = \left(\frac{2}{13}, \frac{3}{13}, \frac{10}{39}, \frac{14}{39} \right)$$

لذا در طولانی مدت احتمال ورود به خانه های ۱ و ۲ و ۳ و ۴ به ترتیب $\frac{2}{13}$ و $\frac{3}{13}$ و $\frac{10}{39}$ و $\frac{14}{39}$ است

حالت های جاذب (Absorbing States)

تعریف ۶ - حالت a_i از زنجیره مارکوف را جاذب نامند، اگر سیستم در وضعیت a_i باقی بماند. بنابراین حالت a_i جاذب است اگر فقط اگر نا این سطر ماتریس انتقال P دارای عنصر ۱ در روی قطر اصلی و عناصر ۰ در سایر جاها باشد (توجه: قطر اصلی ماتریس $A = (a_{ij})_{n \times n}$ عبارتست از درایه های $a_{11}, a_{22}, a_{33}, \dots, a_{nn}$)

سوال ۱۰ - فرض کنید ماتریس زیر ماتریس انتقال یک زنجیره مارکوف است

$$P = \begin{matrix} & \begin{matrix} a_1 & a_2 & a_3 & a_4 & a_5 \end{matrix} \\ \begin{matrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{matrix} & \begin{pmatrix} \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

حالت a_5 و a_2 جاذب اند

سوال ۱۱ - همان حرکت تصادفی روی محور x روی نقاط $\{1, 2, 3, 4, 5\}$ (سوال ۳) با این تفاوت در نظر می گیریم که اگر تکرار به نقاط ۰ و یا ۵ می رسد، در گام بعدی در همان نقطه متوقف می شود و:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 \\ 0 & q & 0 & p & 0 \\ 0 & 0 & q & 0 & p \\ 0 & 0 & 0 & q & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

قضیه ۹ - اگر ماتریس P در یک زنجیره مارکوف دارای ۱ در قطر اصلی باشد
آنگاه P منظم نخواهد بود. (به جز اینکه ماتریس P ، 1×1 باشد)

اثبات - با استقراء ثابت می شود که هر توانی $(P^n: n \in \mathbb{N})$ از P هم دارای ۱ در قطر اصلی است.

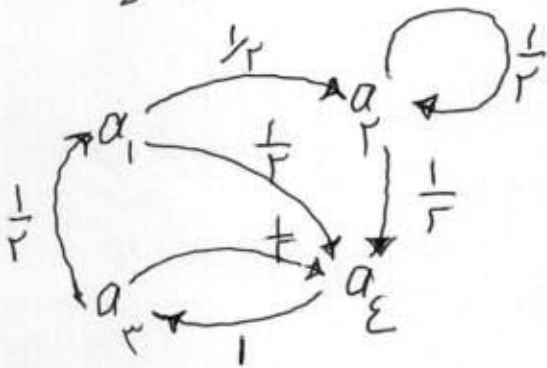
دیاگرام انتقال

یکی از روشهای نمایش ماتریس انتقال، دیاگرام انتقال (transition diagram) است. آن یک گراف جهت دار است که رئوس گراف نمایش حالت که وصال $q_i \rightarrow q_j$ با احتمال p_{ij} مشخص شده است، در حقیقت وصال $q_i \rightarrow q_j$ وجود دارد، اگر $p_{ij} > 0$.

$$P = \begin{matrix} & \begin{matrix} a_1 & a_2 & a_3 & a_4 \end{matrix} \\ \begin{matrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{matrix} & \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix}$$

مثال: تناظر با ماتریس انتقال:

گراف جهت دار زیر است:

مسائل برای حل

۱- شخصی اگر در هر یک از نقاط A ، B ، C ، D قرار گیرد با احتمال $\frac{1}{4}$ به هر یک از نقاط مجاور می رود. اگر اکنون در نقطه A باشد، پس از ۵ بار حرکت با چه احتمالی در نقطه B قرار می گیرد؟

جواب $\frac{1}{4}$

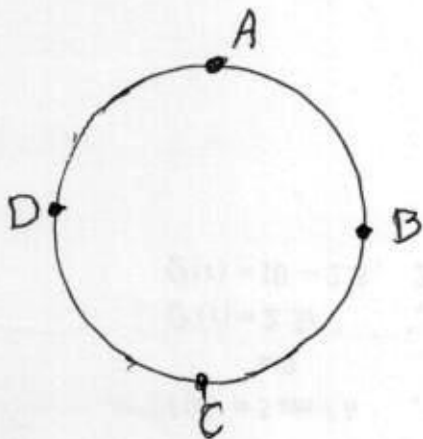
۲- نشان دهید ماتریس انتقال P منظم است.

$$P = \begin{pmatrix} 0 & 0 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{pmatrix}$$

P^n وقتی که $n \rightarrow \infty$ به سمت چه ماتریسی میل می کند؟

جواب: تمام عناصر $\frac{2}{5}$ مثبت اند.

$$\lim_{n \rightarrow \infty} P^n = \begin{pmatrix} \frac{2}{5} & \frac{2}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{2}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{2}{5} & \frac{2}{5} \end{pmatrix}$$



۳- مطلوب است α و β و γ بطوریکه P یک ماتریس استوکاستیک باشد

$$P = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 1 & 2 \\ 0 & 3 & 4 \end{pmatrix}$$

۴- دستگاه فرضی به تصادف سیگنال‌های ۰ و ۱، ارسال می‌دارد، بطوریکه با احتمال α دو سیگنال متوالی یکسان و با احتمال β دو سیگنال متوالی متفاوت ارسال می‌شود و $\alpha + \beta = 1$ ، ماتریس انتقال این فرایند را بنویسید و نشان دهید:

$$P^n = \begin{pmatrix} \frac{1 + (\alpha - \beta)^n}{2} & \frac{1 - (\alpha - \beta)^n}{2} \\ \frac{1 - (\alpha - \beta)^n}{2} & \frac{1 + (\alpha - \beta)^n}{2} \end{pmatrix}$$

و به دو طریق دو طولانی مدت فراوانی سیگنال‌های ۰ و ۱، راحت بکنید.

۵- اگر $0 < a, b < 1$ ، نشان دهید:

$$\lim_{n \rightarrow \infty} \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}^n = \begin{pmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{b}{a+b} & \frac{a}{a+b} \end{pmatrix}$$

۶- با فرض:

$$P = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

نظم است، P^n وقتی که $n \rightarrow \infty$

به یک ماتریس میل می‌کند؟

۷- نشان دهید ماتریس استوکاستیک $A = \begin{pmatrix} 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & 0 \end{pmatrix}$ ، نظم است

۸- ماتریس استوکاستیک $P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 1 & 0 \end{pmatrix}$ و توزیع احتمالات اولیه $P^{(0)} = (\frac{1}{3}, 0, \frac{2}{3})$ مفروض اند. مطلوب است محاسبه:

$$\Rightarrow 1) P_{2,3}^{(2)}, 2) P_{1,3}^{(2)}, 3) P_{1,3}^{(4)}, 4) P_{1,3}^{(4)}, 5) \lim_{n \rightarrow \infty} P^{(0)} P^n, 6) \lim_{n \rightarrow \infty} P^n$$

جواب‌ها:

$$\Rightarrow P_{2,3}^{(2)} = \frac{1}{2}, P_{1,3}^{(2)} = 0, P_{1,3}^{(4)} = (\frac{1}{4}, \frac{1}{12}, \frac{1}{6}), P_{1,3}^{(4)} = \frac{1}{6}, \lim_{n \rightarrow \infty} P^{(0)} P^n = (\frac{2}{3}, \frac{1}{3}, \frac{1}{3})$$

۹- یک تاس سالم را متوالیاً n بار می‌کنیم. فرض کنید X_n ما کس عدد رو شده در n بار تاس متوالی باشد، فضای حالت و ماتریس استوکاستیک این فرایند را بنویسید. آیا این ماتریس منتظم است؟ - ۱۰
 به اولین n بار تاس را بنویسید و سپس $P^{(2)}$ و $P^{(3)}$ را حساب کنید.
 جواب - فضای حالت $\{1, 2, 3, 4, 5, 6\}$ ، ماتریس استوکاستیک P

$$P = \begin{pmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ 0 & \frac{5}{6} & \frac{1}{6} & 0 & 0 & 0 \\ 0 & 0 & \frac{5}{6} & \frac{1}{6} & 0 & 0 \\ 0 & 0 & 0 & \frac{5}{6} & \frac{1}{6} & 0 \\ 0 & 0 & 0 & 0 & \frac{5}{6} & \frac{1}{6} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$P^{(1)} = \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6} \right)$$

$$P^{(2)} = \left(\frac{1}{36}, \frac{5}{36}, \frac{1}{36}, \frac{1}{36}, \frac{1}{36}, \frac{11}{36} \right)$$

$$P^{(3)} = \left(\frac{1}{216}, \frac{7}{216}, \frac{19}{216}, \frac{37}{216}, \frac{61}{216}, \frac{91}{216} \right)$$

P منتظم نیست.

۱- نشان دهید که ماتریس استوکاستیک :

$$P = \begin{pmatrix} 0 & 0 & p & q \\ 0 & 0 & q & p \\ p & q & 0 & 0 \\ q & p & 0 & 0 \end{pmatrix} \quad (p+q=1)$$

منتظم نیست. اگر $\begin{pmatrix} p & q \\ q & p \end{pmatrix} = R$ باشد داریم :

$$P = \begin{pmatrix} R^{2n} & 0 \\ 0 & R^{2n} \end{pmatrix} \quad \text{و} \quad P^{2n-1} = \begin{pmatrix} 0 & R^{2n-1} \\ R^{2n-1} & 0 \end{pmatrix} \quad \text{که} \quad 0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

۱۱- قضیه ۶ را به ازای $n=2$ ثابت کنید.

۱۲- اثبات قضیه ۷.

۱۳- جعبه A شامل ۲ مهره سفید و جعبه B شامل ۴ مهره قرمز است. هر بار به تصادف یک مهره از هر جعبه را انتخاب کرده و به جایی که می‌خواهید آنجا می‌گذارید. فرض کنید که X_n تعداد مهره که قرمز در جعبه A پس از n گام باشد.

۱- ماتریس استوکاستیک P این فرایند را بنویسید.

۲- مطلوب است احتمال اینکه پس از سه گام ۲ مهره قرمز در A باشد.

۳- مطلوبست انتقال لیگه در طولانی مرحله جاگزینی ۲ نره قرمز در A باشد.

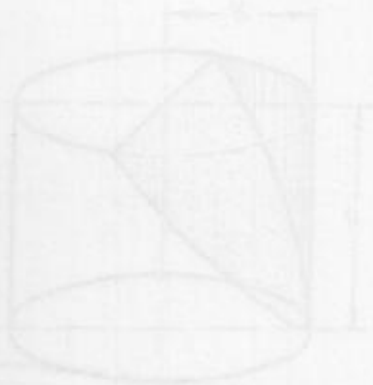
جواب ها:

$$1- \quad P = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{8} & \frac{1}{2} & \frac{3}{8} \\ 0 & \frac{1}{4} & \frac{1}{4} \end{pmatrix}, \quad 2- \frac{3}{8}, \quad 3- \frac{2}{5}$$

۴- سه نره قبل را وقتی که در آغاز ۳ نره سفید در جعبه A و ۳ نره قرمز در جعبه قرمز باشد حل کنید.

$$\text{جواب ها: ۱- ماتریس انتقال} \quad P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{9} & \frac{4}{9} & \frac{4}{9} & 0 \\ 0 & \frac{4}{9} & \frac{4}{9} & \frac{1}{9} \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$2- \frac{32}{81}, \quad 3- \frac{9}{20}$$



زنجیره مارکوف (صحنه ۱) - ۱۷

اثبات قضیه ۶ به ازای $n=2$

قضیه ۶ - اگر P ماتریس انتقال یک زنجیره مارکوف باشد، در اینصورت ماتریس

انتقال n -گام (که $P^{(n)}$ نشان می‌دهیم) برابر است با $P^{(n)} = P^n$

اثبات - برای اثبات $P^{(2)} = P^2$ کافیت ثابت کنیم به ازای هر i, j

در $m=1, 2, 3, \dots$ عنصر i -ام ماتریس $P^{(2)}$ برابر است با عنصر i -ام ماتریس P^2

ولی عنصر i -ام ماتریس $P^{(2)}$ عبارتست از احتمال تغییر وضعیت سیستم از q_i به q_j پس از دو گام یعنی:

$$\begin{aligned} P_{ij}^{(2)} &= P(q_i \rightarrow q_1 \rightarrow q_j \text{ یا } q_i \rightarrow q_2 \rightarrow q_j \text{ یا } \dots \text{ یا } q_i \rightarrow q_m \rightarrow q_j) \\ &= \sum_{k=1}^m P(q_i \rightarrow q_k \rightarrow q_j) = \sum_{k=1}^m P(q_i \rightarrow q_k) P(q_k \rightarrow q_j) = \sum_{k=1}^m P_{ik} P_{kj} \end{aligned}$$

عنصر i -ام ماتریس P^2 = عنصر i -ام حاصلضرب ماتریس P در ماتریس P

اثبات قضیه ۷

قضیه ۷ - اگر P ماتریس انتقال یک زنجیره مارکوف و $p = (p_1, p_2, \dots, p_m)$

توزیع احتمال سیستم در لحظه ای دلخواه باشد، آنگاه pP توزیع احتمال سیستم در گام بعدی و pP^n توزیع احتمال سیستم در n گام بعدی است.

اثبات: کافیت نشاندن دهم احتمال اینکه در گام بعدی q_j باشد برابر است با عنصر i -ام ماتریس pP اگر این احتمال را به $P(q_j)$ نشاندیم داریم:

(احتمال اینکه سیستم از وضعیت q_i به وضعیت q_j تغییر کند) $\times$ (احتمال اینکه سیستم در لحظه آغازی در وضعیت q_k باشد)

$$P(q_j) = \sum_{k=1}^m p_k P(q_k \rightarrow q_j)$$

حاصلضرب ماتریس p در P = $\sum_{k=1}^m p_k P(q_k \rightarrow q_j) = \sum_{k=1}^m p_k p_{kj}$

عنصر i -ام ماتریس pP

پس همین بدول توزیع احتمال سیستم در دو گام بعدی $(pP) \times (pP)$ یعنی pP^2 و ...

توزیع نرمال - ۱

۱- توزیع نرمال معروف ترین توزیع پیوسته است، که در آن دامنه متغیر از $-\infty$ تا $+\infty$ است. تابع چگالی احتمال این توزیع:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\beta} e^{-\frac{(x-\alpha)^2}{2\beta^2}} \quad \beta > 0 \quad (1)$$

است. ضریب $\frac{1}{\sqrt{2\pi}\beta}$ از این جهت لحاظ شده است. زیرا

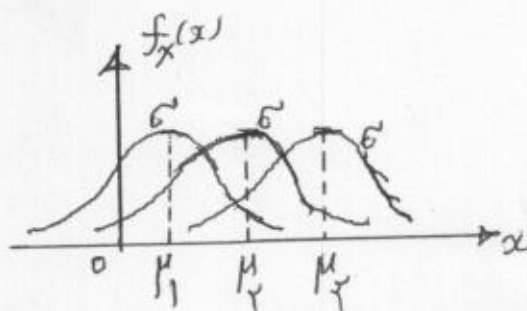
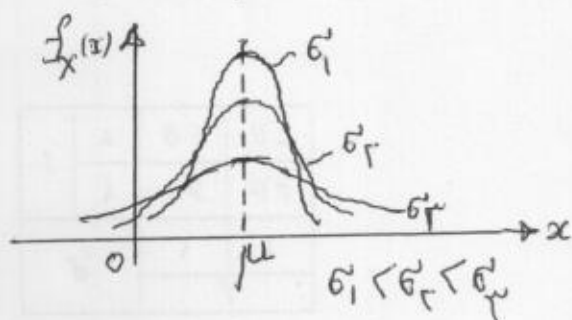
$$\int_{-\infty}^{\infty} e^{-\frac{(x-\alpha)^2}{2\beta^2}} dx = \sqrt{2\pi}\beta$$

یا گوییم این توزیع $\mu = E(X) = \alpha$ و واریانس توزیع:

$$\text{Var}(X) = \sigma^2 = E(X - \mu)^2 = \beta^2$$

از این جهت عملاً از حروف μ و σ به عنوان پارامتر در تابع چگالی توزیع استفاده می شود.

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty$$



معمولاً تابع $f_X(x)$ دارای محور تقارن $x = \mu$ است بنابراین اگر این توزیع صفر است. μ را پارامتر مکان (location) (scale) می نامند.

را با نماد $N(\mu, \sigma^2)$ نشان می دهند

$$X \sim N(\mu, \sigma^2)$$

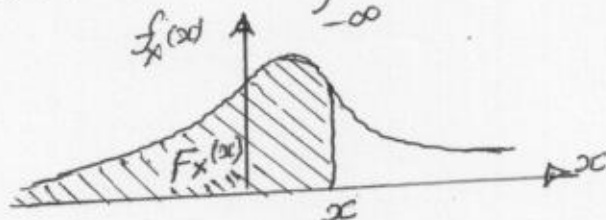
دامنه متغیر از $-\infty$ تا $+\infty$ متغیر می گنند
از سه برابر انحراف معیار توزیع (یعنی σ) بیشتر باشد

توزیع نرمال ۲-
 عملاً آن را به عنوان توزیعی که اندازه متغیر شیب باشد، در نظر گرفته می شود. توزیع نرمال با میانگین $\mu=0$ و $\sigma^2=1$ را توزیع نرمال استاندارد گفته می شود. و برای آن جدول سطح زیر منحنی تابع چگالی احتمال که معروف به فرادانی منبى و احتمال است تنظیم شده است.

$$f_z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad (1)$$

۲- تابع تجمعی توزیع نرمال :

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$



و با یکجا رفتن تغییر متغیر $\frac{t-\mu}{\sigma} = u$:

$$P(X \leq x) = \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

و این به دست سطح زیر منحنی توزیع نرمال $N(0,1)$ است. معمولاً در برخی از کتب آماری با توجه به تفاوت نمودار تابع چگالی توزیع نرمال این مساحت از سیدات نامبر 360 جدول شده است.

فرمول آبی تقریبی هم برای محاسبه $P(X \leq x)$ بدست آمده است از جمله فرمول زیر (استخرج از کتاب معروف آبراموویتز (Milton Abramowitz)

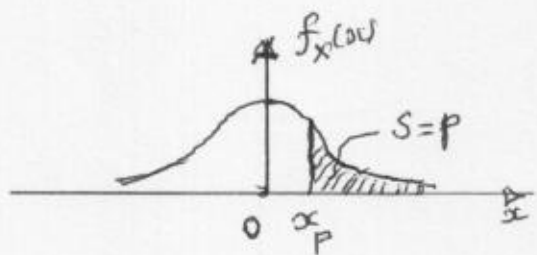
$$P(X \leq x) = P\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = P(Z_{N(0,1)} \leq t) \quad \text{و} \quad t = \frac{x-\mu}{\sigma}$$

$$= 1 - \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} (a_1 q + a_2 q^2 + a_3 q^3) + \epsilon(t) \quad ; \quad q = \frac{1}{1+t^2}$$

$$a = 0.332947, \quad a_1 = 0.473118, \quad a_2 = -0.12017, \quad a_3 = 0.93722, \quad \epsilon(t) < 1 \times 10^{-5}$$

توزیع نرمال ۳-

و اگر x_p چنان باشد که مساحت سطح زیر منحنی توزیع (ادویه) از x_p با بینهایت p باشد و $0 < p < 1$.



$$\left(\frac{1}{\sqrt{2\pi}} \int_{x_p}^{\infty} e^{-\frac{t^2}{2}} dt = p \right)$$

آنگاه داریم:

$$x_p = m - \frac{a_0 + a_1 m}{1 + b_1 m + b_2 m^2} + \varepsilon(p)$$

که در این فرمول $m = \sqrt{\ln \frac{1}{p}}$ و $a_0 = 2.30753$, $a_1 = 0.27061$ و $b_1 = 0.99229$, $b_2 = 0.4481$; $|\varepsilon(p)| < 3 \times 10^{-3}$

یا استفاده از توزیع (ادویه) می توان نشان داد 98.2% افراد جامعه نرمال در فاصله $\mu - \sigma$ تا $\mu + \sigma$ ، 95.4% در فاصله $\mu - 2\sigma$ تا $\mu + 2\sigma$ و 99.74% در فاصله $\mu - 3\sigma$ تا $\mu + 3\sigma$ قرار دارد و فقط 0.2% افراد جامعه در خارج از بازه $[\mu - 3\sigma, \mu + 3\sigma]$ واقع می باشد، یعنی احتمال آنگاه اندازه فرکانس از جامعه $N(\mu, \sigma)$ کمتر از $\mu - 3\sigma$ باشد، 0.13% است.

۳- خواص توزیع نرمال:

(۱) - اگر x دارای توزیع $N(\mu, \sigma)$ باشد، آنگاه متغیر $y = a + bx$

که a و b تغییر ناپذیر هستند دارای توزیع $N(a + b\mu, b^2\sigma^2)$ است. در حالت خاص متغیر $z = \frac{x - \mu}{\sigma}$ دارای توزیع (ادویه) است. تقسیم این خاصیت از این قرار است. اگر متغیرهای x_i ($i=1, 2, \dots, n$) داده اند، بطوریکه $x_i \sim N(\mu_i, \sigma_i^2)$

آنگاه متغیر $y = a + b_1 x_1 + b_2 x_2 + \dots + b_n x_n$ دارای توزیع $N(a + b_1 \mu_1 + b_2 \mu_2 + \dots + b_n \mu_n, b_1^2 \sigma_1^2 + b_2^2 \sigma_2^2 + \dots + b_n^2 \sigma_n^2)$ است.

توزیع نرمال - ۴

(۲) اگر $x_1, x_2, \dots, x_n$ نمونه‌های استحصای از یک جامعه $N(\mu, \sigma^2)$ باشد
آنگاه میانگین نمونه‌ها: $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ (که متغیریت تصادفی) دارای
توزیع $N(\mu, \frac{\sigma^2}{n})$ و متغیر: $z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$ دارای توزیع $N(0, 1)$ است.

(۳) - برای یک توزیع آماری دلخواه که گشتاور مرتبه دوم (یا واریانس) آن محدود
باشد، متغیر $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ برای n کی بزرگ دارای توزیع مجانبی
 $N(\mu, \frac{\sigma^2}{n})$ است، که μ و σ^2 به ترتیب میانگین و واریانس آن

توزیع است (قضیه حد مرکزی). حتی این قضیه وقتی که توزیع جامعه
ناپوسته باشد (مثل توزیع دوجمله‌ای و توزیع پواسن) برقرار است.

(۴) کاربرد اصلی توزیع نرمال: توزیع نرمال برای تعیین حدود اعتماد
پارامتری برآورد شده از یک جامعه، همچنین آزمون فرض پارامترهای

توزیع های پارامتری (قضیه حد مرکزی) و محاسب احتمال برای
توزیع هائی که دارای پارامتری با اندازه بزرگ باشد، بکار می رود.

توزیع های مشتق شده از توزیع نرمال

۱- توزیع کای-دو که متغیر آن با نماد χ_n^2 نشان داده می شود

عبارت است از مجموع مربعات n متغیر نرمال $N(0, 1)$ که مستقل
از یکدیگر توزیع شده اند.

یعنی: $\chi_n^2 = \sum_{i=1}^n x_i^2$ که $x_i \sim N(0, 1)$ تابع چگالی این توزیع:

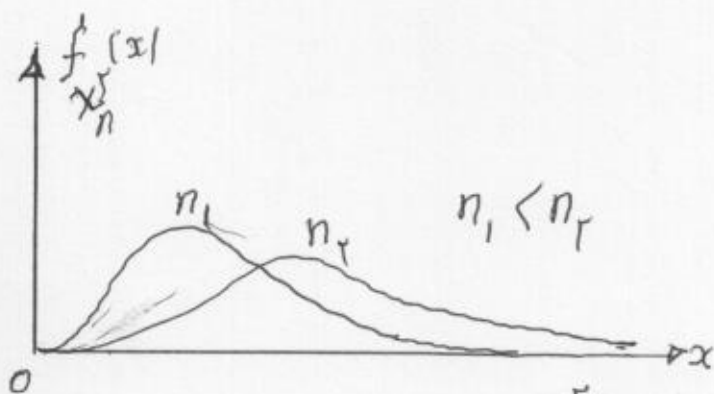
$$f_{\chi_n^2}(x) = \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} e^{-\frac{x}{2}} x^{\frac{n}{2}-1} \quad x > 0$$

و این یک توزیع گاما، با پارامترهای $\gamma = \frac{n}{2}$ و $\theta = \frac{1}{2}$ است.

میانگین و واریانس این توزیع به ترتیب $\mu = n$ و $\sigma^2 = 2n$ است.

توزیع‌های مشتق شده از توزیع نرمال - ۵

این توزیع چوله به راست است (ضریب چولگی آن مثبت می‌باشد).
 و برای n بزرگتر شود تابع چگالی این توزیع به سمت متقارن شدن
 گرایش دارد. برای $n=2$ ، توزیع متغیر X_2 توزیع نمایی:
 $f(x) = \frac{1}{2} e^{-\frac{x}{2}}$ است. ثابت می‌شود برای مقادیر بزرگ
 n ، این توزیع به سمت توزیع $N(n, n)$ گرایش دارد و متغیر
 $Z = \frac{X_n^2 - n}{\sqrt{2n}}$ برای n بزرگ دارای توزیع جانبی $N(0, 1)$ است.



توجه: اگر از یک جامعه $N(\mu, \sigma^2)$ نمونه تصادفی $X_1, X_2, \dots, X_n$ انتخاب کنیم متغیر:

$$U = \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma} \right)^2 = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2}$$

توزیع χ_n^2 است. ولی اگر در محاسبه U بجای μ (که میانگین واقعی جامعه است)، میانگین نمونه هاستی
 آنگاه متغیر $V = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2}$ قرار دهیم

دارای توزیع χ_{n-1}^2 است زیرا متغیرهای $X_1 - \bar{X}, X_2 - \bar{X}, \dots, X_n - \bar{X}$

چون مجموع آن‌ها مساوی صفر است. لذا این صفت به اندیس
 n در توزیع χ_n^2 درم آزادگی توزیع گفته می‌شود.

توزیع‌های مشتق شده از توزیع نرمال - ۶

و وقتی که $n \rightarrow \infty$ توزیع متغیر x_n دارای توزیع مجانبی $N(n, 2n)$ و متغیر $U = \frac{x_n^2 - n}{\sqrt{2n}}$ دارای توزیع مجانبی $N(0, 1)$ است. (البته نسبت شده به متغیر $\sqrt{2} x_n^2$ سریع تر به سمت توزیع $N(\sqrt{2n}, 1)$ میل می‌کند).

یکی از کاربردهای توزیع گامی - دو برآزش یک توزیع مناسب روی یک سری داده‌ای آماری است (آزمون نیکویی برآزش)

۲- توزیع t - استودنت :

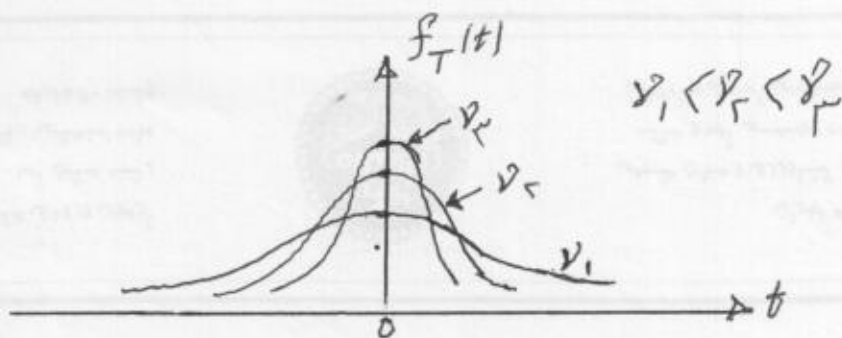
دیدیم در توزیع نرمال $N(\mu, \sigma^2)$ متغیر تصادفی $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ که

$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$ و $x_i \sim N(\mu, \sigma^2)$ ($i=1, 2, \dots, n$) دارای توزیع $N(0, 1)$ می‌باشد، حال اگر واریانس جامعه (که) را ندانیم و بجای آن از برآورد آن یعنی $s^2 = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n-1}$ استفاده ده کنیم، دیگر متغیر تصادفی

$T = \frac{\bar{X} - \mu}{s/\sqrt{n}}$ دارای توزیع نرمال $N(0, 1)$ نیست. در آمار ریاضی نسبت می‌شود که متغیر T دارای توزیعی با تابع چگالی احتمال:

$$f_T(t) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi\nu} \Gamma(\frac{\nu}{2})} \frac{1}{(1 + \frac{t^2}{\nu})^{\frac{\nu+1}{2}}} \quad -\infty < t < \infty$$

که در آن $\nu = n-1$. میانگین این توزیع $E(T) = \mu = 0$ و واریانس آن $Var(T) = \frac{\nu}{\nu-2}$ است، این توزیع مانند توزیع نرمال متقارن است، بنابراین فریب چولگی آن صفر است.



توزیع آبی مشتق شده از توزیع نرمال - v

وقتی که $v \rightarrow \infty$ ، توزیع t-استودنت به سمت توزیع $N(0,1)$ میل می کند

یعنی: $\lim_{v \rightarrow \infty} \frac{\Gamma(\frac{v+1}{2})}{\sqrt{\pi v} \Gamma(\frac{v}{2})} \frac{1}{(1 + \frac{t^2}{v})^{\frac{v+1}{2}}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$

برای اثبات این فرمول، از فرمول جابجایی استرلینگ برای تابع گاما استفاده می کنیم.

$\Gamma(n) \approx \sqrt{2\pi} e^{-n} n^{n-\frac{1}{2}}$ $n \rightarrow \infty$ لذا داریم:

$$\begin{aligned} f_T(t) &= \frac{\Gamma(\frac{v+1}{2})}{\sqrt{\pi v} \Gamma(\frac{v}{2})} \frac{1}{(1 + \frac{t^2}{v})^{\frac{v+1}{2}}} \approx \frac{\sqrt{2\pi} e^{-\frac{v+1}{2}} (\frac{v+1}{2})^{\frac{v+1}{2}-\frac{1}{2}}}{\sqrt{\pi v} \sqrt{2\pi} e^{-\frac{v}{2}} (\frac{v}{2})^{\frac{v}{2}-\frac{1}{2}}} \times \frac{1}{(1 + \frac{t^2}{v})^{\frac{v+1}{2}}} \\ &= \frac{e^{-\frac{1}{2}}}{\sqrt{\pi v}} \times \frac{(\frac{v+1}{2})^{\frac{v}{2}}}{(\frac{v}{2})^{\frac{v}{2}-\frac{1}{2}}} \times \frac{1}{(1 + \frac{t^2}{v})^{\frac{v+1}{2}}} = \frac{e^{-\frac{1}{2}}}{\sqrt{\pi v} (\frac{v}{2})^{\frac{1}{2}}} \times \left(\frac{v+1}{v}\right)^{\frac{v}{2}} \times \frac{1}{(1 + \frac{t^2}{v})^{\frac{v+1}{2}}} \\ &= \frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}} \times \left(1 + \frac{1}{v}\right)^{\frac{v}{2}} \times \frac{1}{(1 + \frac{t^2}{v})^{\frac{v}{2}}} \times \frac{1}{(1 + \frac{t^2}{v})^{\frac{1}{2}}} \end{aligned}$$

و: $\lim_{v \rightarrow \infty} \left(1 + \frac{1}{v}\right)^{\frac{v}{2}} = e^{\frac{1}{2}}$, $\lim_{v \rightarrow \infty} \left(1 + \frac{t^2}{v}\right)^{\frac{v}{2}} = e^{\frac{t^2}{2}}$

و $\lim_{v \rightarrow \infty} \left(1 + \frac{t^2}{v}\right)^{\frac{1}{2}} = 1$ (توجه: $\lim_{x \rightarrow \infty} \left(1 + \frac{a}{x}\right)^{bx} = e^{ab}$)

لذا: $\lim_{v \rightarrow \infty} f_T(t) = \frac{e^{-\frac{1}{2}}}{\sqrt{2\pi}} \times e^{\frac{1}{2}} \times \frac{1}{e^{\frac{t^2}{2}}} \times \frac{1}{1} = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$

توجه: اگر در توزیع t-استودنت $v > 25$ باشد، با تقریب خوبی آن را توزیع $N(0,1)$ در نظر می گیریم، بنابراین در کتب آمار جدول مربوط

سطح زیر منحنی تابع چگالی احتمال توزیع t-استودنت تا $v=25$ تنظیم شده است این مطلب که برای $v > 25$ از جدول توزیع نرمال استفاده می کنیم حتی اگر جامعه مادر دارای توزیع نرمال نباشد، برابر با قضیه حد مرکزی هم عمل می شود.

توزیع بی مستقیم شده از توزیع نرمال - ۸

از توزیع t - استودنت برای محاسبه حدود اطمینان میانگین جامعه نرمال وقتی که واریانس آن معلوم نباشد و بجای آن برآورد s^2 برای آن اندازه نمونه ها اعمال کنیم و $n \geq 25$ باشد، استفاده ده می شود.

۳- توزیع F (یا توزیع فیشر)

فرض کنیم n_1 متغیر

$x_1, x_2, \dots, x_{n_1}$ هر یک دارای توزیع (اوا) $N(0, 1)$

و n_2 متغیر

$y_1, y_2, \dots, y_{n_2}$ هم دارای توزیع (اوا) $N(0, 1)$ باشند

آنگاه متغیرهای $x_1^2, x_2^2, \dots, x_{n_1}^2$ و $y_1^2, y_2^2, \dots, y_{n_2}^2$ دارای توزیع

طبیعی ترتیب متغیر $(x_1^2)/n_1$ و n_1 آزادی می باشند.

بر آن توزیع $F = \frac{(x_1^2)/n_1}{(y_1^2)/n_2}$ دارای توزیع است که

توزیع دارای تابع چگالی (احتمال):

$$f_{n_1, n_2}(x) = \frac{n_1^{\frac{n_1}{2}} n_2^{\frac{n_2}{2}}}{B(\frac{n_1}{2}, \frac{n_2}{2})} \frac{x^{\frac{n_1}{2}-1}}{(n_2 + n_1 x)^{\frac{n_1+n_2}{2}}} \quad x > 0$$

که در این فرمول $B(m, n)$ تابع بتا است که توسط انگرال

$$B(m, n) = \int_0^1 t^{m-1} (1-t)^{n-1} dt \quad m, n > 0$$

از توزیع F برای آزمون مساوی بودن میانگین چند جامعه (تجزیه واریانس) استفاده ده می شود.

میانگین و واریانس توزیع F به ترتیب

$$\mu = \frac{n_2}{n_2 - 2}$$

$$s^2 = \frac{2n_2^2(n_1 + n_2 - 2)}{n_1(n_2 - 2)^2(n_2 - 4)}$$

توزیع های مشتق شده از توزیع نرمال - ۹

۷- توزیع گول نرمال (Lognormal Distribution)

تولید - اگر تغییر x در دارای توزیع نرمال باشد، گفته می شود تغییر x دارای توزیع لوگ نرمال است. پس به فرض $y = \ln x$ ، $x = e^y$ ، یعنی اگر y دارای توزیع نرمال باشد، e^y دارای توزیع لوگ نرمال است. لذا داده تغییرات x تعدادی نسبت است.

فرض کنید (۱) $f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_Y} e^{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}}$ و $f_X(x)$ به ترتیب تابع چگالی احتمال متغیرهای Y و X باشند و متناظر با متغیری $Y = y + dy$ و $X = x + dx$ را بنویسید.

$$f_X(x) dx = P(x \leq X \leq x + dx) = P(y \leq Y \leq y + dy) = f_Y(y) dy$$

$$= \frac{1}{\sqrt{2\pi}\sigma_Y} e^{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}} dy \Rightarrow f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_Y} e^{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}} \times \frac{dy}{dx}$$

ولی $y = \ln x$ و $\frac{dy}{dx} = \frac{1}{x}$ ، لذا تابع کلی اهمیت توزیع کوکس-بول
چنین است:

$$f_X(x) = \frac{1}{\sqrt{\pi}\sigma_y} e^{-\frac{(\ln x - \mu_y)^2}{2\sigma_y^2}} \times \frac{1}{x} \quad (r) \quad x > 0 \quad (r)$$

تدبر کنید μ و σ به عنوان پارامترهای این توزیع محسوب می‌شوند
و این دو به ترتیب میانگین و واریانس توزیع نرمال است، که توزیع گوی نرمال
از آن مشتق می‌شود.

توزیع های مشتق شده از توزیع نرمال - ۱۰ -

معمولاً بیکسبی و وارپاش توزیع لوگ نرمال

$$\mu_x = E(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \frac{1}{\sqrt{2\pi}\sigma_y} \int_{-\infty}^{\infty} e^{-\frac{(\ln x - \mu_y)^2}{2\sigma_y^2}} dx$$

در این کسری تغییر متغیر $\ln x = y$ یعنی $x = e^y$ بکار می بریم: $dx = e^y dy$

$$\text{لذا: } \begin{cases} x \rightarrow 0^+ \Rightarrow y \rightarrow -\infty \\ x \rightarrow \infty \Rightarrow y \rightarrow +\infty \end{cases}$$

$$\mu_x = \frac{1}{\sqrt{2\pi}\sigma_y} \int_{-\infty}^{\infty} e^{-\frac{(y - \mu_y)^2}{2\sigma_y^2}} + y dy$$

ولی:

$$-\frac{(y - \mu_y)^2}{2\sigma_y^2} + y = -\frac{1}{2\sigma_y^2} (y^2 - 2\mu_y y + \mu_y^2 - 2\sigma_y^2 y)$$

$$= -\frac{1}{2\sigma_y^2} \{ y^2 - 2(\mu_y + \sigma_y^2) + \mu_y^2 \} = -\frac{1}{2\sigma_y^2} \{ (y - (\mu_y + \sigma_y^2))^2 - (\mu_y + \sigma_y^2)^2 + \mu_y^2 \}$$

$$= -\frac{1}{2\sigma_y^2} \{ (y - (\mu_y + \sigma_y^2))^2 - 2\mu_y \sigma_y^2 - \sigma_y^4 \} = -\frac{(y - (\mu_y + \sigma_y^2))^2}{2\sigma_y^2} + \mu_y + \frac{\sigma_y^2}{1}$$

$$\mu_x = e^{\mu_y + \frac{\sigma_y^2}{1}} \left[\frac{1}{\sqrt{2\pi}\sigma_y} \int_{-\infty}^{\infty} e^{-\frac{(y - (\mu_y + \sigma_y^2))^2}{2\sigma_y^2}} dy \right] \quad \text{لذا:}$$

ولی حاصل کسری درون [] مساوی یک است. لذا:

$$\boxed{\mu_x = e^{\mu_y + \frac{\sigma_y^2}{1}}} \quad (3)$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \frac{1}{\sqrt{2\pi}\sigma_y} \int_{-\infty}^{\infty} e^{-\frac{(y - \mu_y)^2}{2\sigma_y^2}} + 2y dy = e^{2\mu_y + 2\sigma_y^2}$$

$$\Rightarrow \sigma_x^2 = E(X^2) - \mu_x^2 = e^{2\mu_y + 2\sigma_y^2} - e^{2\mu_y + \sigma_y^2} = e^{2\mu_y + \sigma_y^2} (e^{\sigma_y^2} - 1)$$

$$\boxed{\sigma_x^2 = \mu_x^2 (e^{\sigma_y^2} - 1)} \quad (4) \quad \text{و یا:}$$

توزیع های مشتق شده از توزیع نرمال - ۱۱

برآورد پارامتریکی توزیع لوگ نرمال :
 همانطور که قبلاً بیان شد μ_y و σ_y^2 به عنوان پارامترهای توزیع لوگ نرمال مطرح اند. فرض کنید $x_1, x_2, \dots, x_n$ یک نمونه گیری به حجم n از جامعه لوگ نرمال باشد. برآورد میانگین و واریانس این توزیع به ترتیب عبارتند از:

$$\hat{\mu}_x = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{و} \quad \hat{\sigma}_x^2 = s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$
 لذا داریم :

$$(3) \Rightarrow \hat{\mu}_x = \bar{x} = e^{\hat{\mu}_y + \frac{\hat{\sigma}_y^2}{2}} \quad (5)$$

$$(4) \Rightarrow \hat{\sigma}_x^2 = s^2 = \bar{x}^2 (e^{\hat{\sigma}_y^2} - 1) \quad (6)$$

از مساوی (۶) نتیجه می شود :

$$e^{\hat{\sigma}_y^2} = \frac{s^2}{\bar{x}^2} + 1 \Rightarrow \hat{\sigma}_y^2 = \ln \left(1 + \frac{s^2}{\bar{x}^2} \right) \quad (7)$$

و با بردن $e^{\hat{\mu}_y}$ در (۵) و استخراج $e^{\hat{\mu}_y}$:

$$\Rightarrow e^{\hat{\mu}_y} = \frac{\bar{x}}{\sqrt{1 + \frac{s^2}{\bar{x}^2}}} \Rightarrow \hat{\mu}_y = \ln \frac{\bar{x}}{\sqrt{1 + \frac{s^2}{\bar{x}^2}}} \quad (8)$$

حساب افعال (رابطه تحشی) در توزیع نرمال :

آمر x دارای توزیع لوگ نرمال با پارامترهای μ_y و σ_y^2 باشد، محاسبه $P(X \leq a)$ منجر به افعال مربوط به توزیع $N(0,1)$ می شود. برای اینکار :

$$P(X \leq a) = P(\ln X \leq \ln a) = P\left(\frac{\ln X - \mu_y}{\sigma_y} \leq \frac{\ln a - \mu_y}{\sigma_y}\right) = \dots$$

$$Z \sim N(0,1) \text{ و در نتیجه: } X = \ln X \sim N(\mu_y, \sigma_y^2) \Rightarrow P\left(Z \leq \frac{\ln a - \mu_y}{\sigma_y}\right)$$

$$P(a \leq X \leq b) = P\left(\frac{\ln a - \mu_y}{\sigma_y} \leq Z \leq \frac{\ln b - \mu_y}{\sigma_y}\right) \quad \text{همین:}$$

و $Z \sim N(0,1)$

توزیع کمی مشتق شده از توزیع نرمال - ۱۲

- مسال ۱- جدول زیر آمار دبی لوج (Peak discharge) سالانه مربوط رودخانه ای از سال ۱۸۹۵ تا ۱۹۹۰ بر حسب cfs (فوت مکعب در ثانیه) می باشد. فرض کنید این داده ها از توزیع لوج نرمال پیروی می کنند.
- الف - میانگین و واریانس این توزیع را برآورد کنید
- ب - فراوانی (نسبی) نظری (مورد انتظار) طبقه سوم را حساب کنید.
- پ - دبی لوج با دوره بازگشت ۱۰۰ ساله را حساب کنید.

صنقه (class)	مركز طبقه x_i	فراوانی نسبی f_i
۲۰,۰۰۰ - ۳۰,۰۰۰	۲۵,۰۰۰	۰/۰۳۰
۳۰,۰۰۰ - ۴۰,۰۰۰	۳۵,۰۰۰	۰/۰۴۵
۴۰,۰۰۰ - ۵۰,۰۰۰	۴۵,۰۰۰	۰/۱۵۲
۵۰,۰۰۰ - ۶۰,۰۰۰	۵۵,۰۰۰	۰/۱۳۶
۶۰,۰۰۰ - ۷۰,۰۰۰	۶۵,۰۰۰	۰/۱۶۷
۷۰,۰۰۰ - ۸۰,۰۰۰	۷۵,۰۰۰	۰/۱۵۲
۸۰,۰۰۰ - ۹۰,۰۰۰	۸۵,۰۰۰	۰/۱۸۲
۹۰,۰۰۰ - ۱۰۰,۰۰۰	۹۵,۰۰۰	۰/۰۹۱
۱۰۰,۰۰۰ - ۱۱۰,۰۰۰	۱۰۵,۰۰۰	۰/۰۰۰
۱۱۰,۰۰۰ - ۱۲۰,۰۰۰	۱۱۵,۰۰۰	۰/۰۴۵

حل: الف - $\bar{X} = \sum_{i=1}^n f_i x_i = 10^3 (0.30 \times 25 + \dots + 0.45 \times 115) = 97500$ cfs

$$s_x^2 = \sum_{i=1}^n f_i (x_i - \bar{X})^2 = 0.30 (25000 - 97500)^2 + \dots + 0.45 (115000 - 97500)^2$$

$$= 441 \times 10^6 \text{ cfs}^2 \Rightarrow s_x = 21000 \text{ cfs}$$

توزیع ریشتن شده از توزیع نرمال - ۱۳

$$\frac{s_x}{\bar{x}} = \frac{27000}{47500} = 0.569 \Rightarrow \hat{\mu}_y = \ln \frac{\bar{x}}{\sqrt{1 + (\frac{s_x}{\bar{x}})^2}} = \ln \frac{47500}{\sqrt{1 + (0.569)^2}}$$

$$\Rightarrow \hat{\mu}_y = 11.0737 \quad \hat{\sigma}_y = \sqrt{\ln \left(\frac{\bar{x}^2}{s_x^2} + 1 \right)} = \sqrt{\ln \left((0.569)^2 + 1 \right)} = 0.95$$

ب- $f'_2 = P(\ln(x_{10}^L) \leq X \leq \ln(x_{10}^U)) = P(\ln(x_{10}^L) \leq \ln Y \leq \ln(x_{10}^U))$

$$= P\left(\frac{\ln(x_{10}^L) - \hat{\mu}_y}{\hat{\sigma}_y} \leq \frac{\ln Y - \hat{\mu}_y}{\hat{\sigma}_y} \leq \frac{\ln(x_{10}^U) - \hat{\mu}_y}{\hat{\sigma}_y} \right)$$

$$= P(-1.5496 \leq Z \leq -0.1854) \quad Z \sim N(0,1)$$

$$= P(0.1854 \leq Z \leq 1.5496) = 0.4418 - 0.2981 = 0.1437$$

پ: اندازه دی کوج با دوره بازگشت ۱۰۰ ساله رابه α شک می دهیم
و این بران معنی اری که:

$$P(X \geq a) = \frac{1}{100}$$

پس داریم: $P(\ln X \geq \ln a) = P\left(\frac{\ln X - \mu_y}{\sigma_y} \geq \frac{\ln a - \mu_y}{\sigma_y} \right) = \frac{1}{100}$

و بی: $z = \frac{\ln X - \mu_y}{\sigma_y} \sim N(0,1)$ و $P(0 \leq z \leq \frac{\ln a - \mu_y}{\sigma_y}) = 0.42$ و از روی

جدول توزیع نرمال

$$\therefore \frac{\ln a - \mu_y}{\sigma_y} = 1.329$$

$$\Rightarrow \ln a = \mu_y + 1.329 \sigma_y = 11.0737 + 1.329 \times 0.95 = 12.1405$$

$$\Rightarrow a = e^{12.1405} = 150965 \approx 150965 \text{ cfs}$$

توزیع‌های مشتق شده از توزیع نرمال - ۱۴

نمونه ۲ - فرض کنید X دارای توزیع دووجهی با پارامترهای:

$$P_X(x) = P(X=x) = \binom{n}{x} p^x q^{n-x}, \quad x=0, 1, 2, \dots, n$$

است. ثابت می‌شود که اگر n به اندازه کافی بزرگ باشد، محاسبه احتمال

$P(a \leq X \leq b)$ را می‌توان با محاسبه $P(a + \frac{1}{2} \leq X \leq b + \frac{1}{2})$ از توزیع:

$$N(\mu = np, \sigma^2 = npq)$$

تقریب زد. فرض کنید $n=25$ و $p=1/3$ باشد و محاسبه $P(4 \leq X \leq 8)$ است. الف - این احتمال از روی توزیع دووجهی:

$$P = P(4 \leq X \leq 8) = \binom{25}{4} \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^{21} + \binom{25}{5} \left(\frac{1}{3}\right)^5 \left(\frac{2}{3}\right)^{20} + \binom{25}{6} \left(\frac{1}{3}\right)^6 \left(\frac{2}{3}\right)^{19} + \binom{25}{7} \left(\frac{1}{3}\right)^7 \left(\frac{2}{3}\right)^{18} + \binom{25}{8} \left(\frac{1}{3}\right)^8 \left(\frac{2}{3}\right)^{17}$$

ب - محاسبه این احتمال از روی توزیع نرمال:

$$\mu = np = 25 \times \frac{1}{3} = 8.33, \quad \sigma^2 = npq = 25 \times \frac{1}{3} \times \frac{2}{3} = 5.56$$

توجه: npq و np به ترتیب همان میانگین و واریانس توزیع دووجهی است.

$$P = P(4 \leq X \leq 8.33) = P\left(\frac{4 - 8.33}{\sqrt{5.56}} \leq Z \leq \frac{8.33 - 8.33}{\sqrt{5.56}}\right)$$

$$= P(-1.873 \leq Z \leq 0) = 0.4086 + 0.1612 = 0.5698$$

که این دو مقدار به هم نزدیک هستند. همین تقریب را می‌توان برای توزیع پواسن با λ بزرگ بکار گرفت.

مسائل

۱- در صورتی بی‌های اوج سالانه یک سوختخانه دارای توزیع گوی نرمال با میانگین

$$\mu_X = 6000 \text{ و انحراف معیار } \sigma_X = 1000$$

باشد. الف - میانگین واریانس

μ_X را حساب کنید. ب - اندازه بی‌های اوج با دوره بازگشت ۲۵ ساله را حساب کنید.

۲- اگر تصحیح‌های: $1, 2, 3, 4, 5, 6, 7, 8, 9, 10$ که به طور مستقل از هم توزیع شده‌اند دارای

توزیع $N(0,1)$ باشند $P(X_1^2 + X_2^2 + X_3^2 + X_4^2 \leq 4)$ را حساب کنید.

۳- میانگین واریانس توزیع‌های ۱- استودنت، ۲- F و ۳- t را حساب کنید.

۴- اگر X دارای توزیع گوی نرمال باشد به طوری که $\ln X \sim N(0,1)$ ، $P(X \leq \mu_X)$ را حساب کنید.

برآورد (Estimate) ۱-

در این قسمت فرض می‌کنیم که اندازه افراد جامعه از توزیع خاصی پیروی می‌کند، مثلاً توزیع نرمال، نمایی، پواسن، گاما، ... عموماً این توزیع‌ها شامل پارامتر ویا پارامترهایی هستند که ممکن است تعریف برخی ویا همه آنها مجهول باشد. منظور از برآورد (Estimate) پارامترهای مجهول یک توزیع عبارتست از تخمین این پارامترها از روی یک نمونه‌گیری $\{x_1, x_2, \dots, x_n\}$ از جامعه وقتی که در محاسبه آنها اصول منطقی و علمی بکار گرفته شود.

اگر به پارامتر توزیع θ برآورد آن را با $\hat{\theta}$ نشان می‌دهیم به طوری که $(x_1, x_2, \dots, x_n) = \hat{\theta}$ ، توهم است که خود یک تفسیر تصادفی است و دارای یک توزیع آماری است.

دور روش اصلی برای پارامتر (یا پارامترهای) یک توزیع آماری بکار گرفته می‌شود

۱- روش گسسته‌ها:

فرض کنید توزیع جامعه دارای ۲ پارامتر α, β و $\alpha, \beta, \dots, \beta$ مجهول باشد، روش گسسته در این قرار است که ابتدا α گسسته در اولین توزیع یعنی

$E(X^j) = \int_0^\infty x^j f_X(x) dx$ ($j=1, 2, \dots$) که وابسته به تعدادی حقیقی ۲ پارامتر هستند، بدست می‌آوریم و پس هر یک از آنها را به وی گسسته‌های هم مرتبه خود از نمونه‌ها، قرار می‌دهیم. یعنی

$$E(X^j) = \sum_{i=1}^n \frac{x_i^j}{n} \quad j=1, 2, 3, \dots, 2 \quad (1)$$

باجل دستگاه (۱) برآورد پارامترها بدست می‌آید.

توهم: اگر جامعه دارای یک پارامتر مجهول باشد، فقط معادله

$\mu(\hat{\alpha}) = E(X) = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$ داریم و در صورتی که دو پارامتر α, β در توزیع مجهول باشد معمولاً از معادلات

$$\mu(\hat{\alpha}, \hat{\beta}) = \bar{x}$$

استفاده می‌کنیم. $\sigma^2(\hat{\alpha}, \hat{\beta}) = s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$

پراورد - ۲

مثال ۱ - برای پراورد دوبارانه α و ν در توزیع گاما با تابع چگالی احتمال

$$f_X(x) = \frac{1}{\alpha^\nu \Gamma(\nu)} e^{-\frac{x}{\alpha}} x^{\nu-1} \quad x > 0, \quad (\alpha > 0, \nu > 0)$$

ابتدا، میانگین و واریانس توزیع که وابسته به α و ν هستند را حساب می‌کنیم

$$\mu = E(X) = \frac{1}{\alpha^\nu \Gamma(\nu)} \int_0^\infty x (e^{-\frac{x}{\alpha}} x^{\nu-1}) dx = \frac{\alpha^{\nu+1} \Gamma(\nu+1)}{\alpha^\nu \Gamma(\nu)} = \alpha \nu$$

$$E(X^2) = \frac{1}{\alpha^\nu \Gamma(\nu)} \int_0^\infty x^2 (e^{-\frac{x}{\alpha}} x^{\nu-1}) dx = \frac{\alpha^{\nu+2} \Gamma(\nu+2)}{\alpha^\nu \Gamma(\nu)} = \alpha^2 \nu (\nu+1)$$

واریانس توزیع :

$$\sigma^2 = E(X^2) - \mu^2 = \alpha^2 \nu$$

حال قرار می‌دهیم :

$$\begin{cases} \hat{\alpha} \hat{\nu} = \hat{\mu} = \bar{X} \\ \hat{\alpha}^2 \hat{\nu} = \hat{\sigma}^2 = S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \end{cases}$$

و از حل این دستگاه درست می‌آوریم :

$$\Rightarrow \hat{\alpha} = \frac{S^2}{\bar{X}} \quad \text{و} \quad \hat{\nu} = \frac{\bar{X}^2}{S^2}$$

۲ - روش حداکثر درست نمایی .

تعریف : اگر $\{X_1, X_2, \dots, X_n\}$ یک نمونه گیری به حجم n از جامعه باشد
اگر جامعه دارای توزیع نامیوسته با تابع احتمال $P_X(x)$ باشد، تابع :

$$L(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P_X(X_i)$$

و اگر جامعه دارای توزیع پیوسته با تابع چگالی احتمال $f_X(x)$ ، تابع :

$$L(X_1, X_2, \dots, X_n) = \prod_{i=1}^n f_X(X_i)$$

را تابع درست نمایی (Likelihood Function) می‌نامند ،

L علاوه بر متغیر نمونه‌ها، تابعی از پارامترهای جامعه است، یعنی

$$L = L(X_1, X_2, \dots, X_n; \alpha_1, \alpha_2, \dots, \alpha_r)$$

برآورد - ۳
روش حداکثر درست نمایی (Maximum Likelihood) برای این است که به داده‌ها و به راجع به x_i ها ($i=1, 2, \dots, n$) چنان برآورد کنیم که تابع L حداکثر شود (یعنی احتمال واقعه‌ای که مشاهده کرده ایم، حداکثر باشد) و شرط لازم برای اینکار

به لحاظ اینکه ما کسب کردن که به تابع نسبت به θ_j $\frac{\partial L}{\partial \theta_j} = 0$ $j=1, 2, 3, \dots, r$
تابع است و از آنجایی که فرم ریاضی بسیاری از توابع گزاشی است. برای سهولت محاسبات از دستگاه معادلات زیر استفاده می‌کنیم

$$\frac{\partial \ln L}{\partial \theta_j} = 0 \quad j=1, 2, 3, \dots, r$$

مثال ۲- برای برآورد پارامتر λ در توزیع پواسن با تابع احتمالی $P_X(x) = e^{-\lambda} \frac{\lambda^x}{x!}$ $x=0, 1, 2, \dots$ و $\lambda > 0$ ، بر اساس نمونه‌گیری: $\{x_1, x_2, \dots, x_n\}$ تابع درست نمایی عبارتست از:

$$L(x_1, x_2, \dots, x_n; \lambda) = \prod_{i=1}^n e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} = e^{-n\lambda} \times \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}$$

$$\ln L = -n\lambda + \left(\sum_{i=1}^n x_i\right) \ln \lambda - \ln\left(\prod_{i=1}^n x_i!\right)$$

$$\frac{\partial (\ln L)}{\partial \lambda} = -n + \left(\sum_{i=1}^n x_i\right) \frac{1}{\lambda} = 0 \Rightarrow \hat{\lambda} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$$

توجه کنید بکارگیری روش گفته درها هم برای λ همین نتیجه را بدست می‌دهد.
مثال ۳- برای جابجایی با توزیع $N(\mu, \sigma^2)$ ابتدا فرض کنید μ مجهول ولی σ^2 معلوم است و پس حالتی را در نظر بگیرید که هر دو پارامتر جابجه مجهول باشد و بخواهیم برآورد روش حداکثر درست نمایی برای برآورد کنیم. برای هر دو روش به تابع درست نمایی:

$$L = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}} \quad \text{برآورد - ۱}$$

$$= \frac{1}{(\sqrt{2\pi})^n} \times \frac{1}{\sigma^n} \times e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2}$$

در حالت اول که فقط μ مجهول است و باید برآورد شود، تنها این معادله داریم

$$\frac{d(\ln L)}{d\mu} = \frac{1}{2\sigma^2} \sum_{i=1}^n 2(x_i - \hat{\mu}) = 0 \Rightarrow \sum_{i=1}^n (x_i - \hat{\mu}) = 0$$

و در حالت دوم که هر دو پارامتر جامعه یعنی μ و σ^2 مجهول باشند:

$$\begin{cases} \frac{\partial(\ln L)}{\partial \mu} = \frac{1}{2\hat{\sigma}^2} \sum_{i=1}^n 2(x_i - \hat{\mu}) = 0 \\ \frac{\partial(\ln L)}{\partial \sigma^2} = -\frac{n}{2\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^4} \sum_{i=1}^n (x_i - \hat{\mu})^2 = 0 \end{cases}$$

و از حل این دستگاه بدست می آوریم:

$$\Rightarrow \hat{\mu} = \bar{X} \quad \text{و} \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \hat{\mu})^2}{n} = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n}$$

در این مثال اگر μ معلوم و σ^2 مجهول که باید برآورد شود: $\hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n}$

تعریف - همانطور که قبلاً ذکر شد $\alpha = \alpha(x_1, \dots, x_n)$ خود یک تغییر تصادفی است

اگر $E(\hat{\alpha}) = \alpha$ باشد، گفته می شود $\hat{\alpha}$ یک برآورد نااریب α است و درستی که $E(\hat{\alpha}) = \alpha + b(\alpha)$ ، $b(\alpha)$ را اریبی این برآورد می نامیم. به عنوان مثال در مثال بالا برای توزیع $N(\mu, \sigma^2)$ وقتی که μ معلوم باشد $\hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n}$ یک برآورد نااریب σ^2 است. و درحالتی که μ و σ^2 مجهول باشند، $\hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n}$ یک برآورد اریب دار σ^2 است و داریم:

$E(\hat{\sigma}^2) = \frac{n-1}{n} \sigma^2 \neq \sigma^2$ ، اینجاست می شود برآورد با یک اریبی روش ماکسیمم درست

دلایل حداقل واریانس، نسبت μ برآورد کننده دیگر است.

توزیع گاما - 1

تعریف - توزیع γ ($\eta \in \mathbb{N}$) متغیر مستقل که هر کدام دارای توزیع نمایی با پارامتر یکسان λ باشد، دارای تابع چگالی احتمال

$$(1) \quad f_X(x) = \frac{\lambda^\eta}{\Gamma(\eta)} x^{\eta-1} e^{-\lambda x} \quad x > 0, \lambda > 0$$

که توزیع گاما نامیده می شود در حالت کلی توزیع گاما برای پارامترهای مثبت η (نه لزوماً صحیح) و λ تعریف می شود، $\Gamma(\eta)$ که تابع گاما نامیده می شود برابر با انتگرال:

$$(2) \quad \Gamma(\eta) = \int_0^\infty t^{\eta-1} e^{-t} dt \quad \eta > 0$$

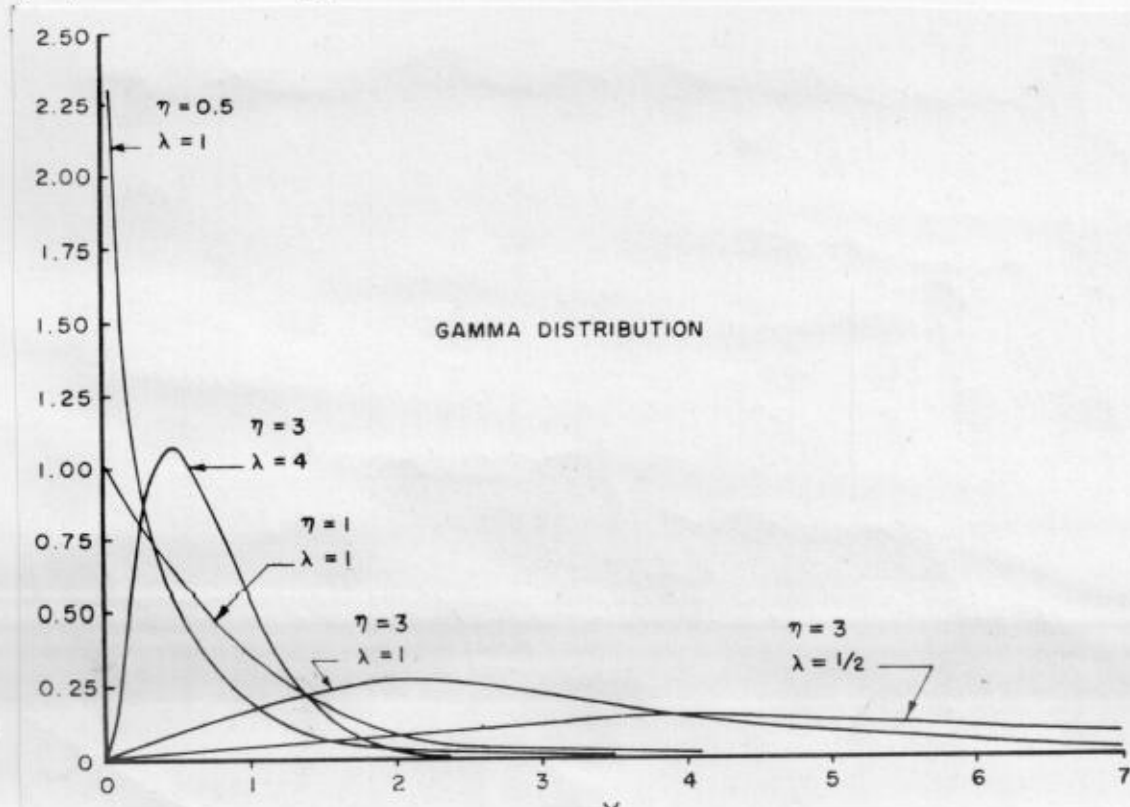
تعریف می شود و برای آن روابط زیر داریم

$$\Gamma(\eta+1) = \eta \Gamma(\eta) \quad \eta > 0 \quad \text{و} \quad \Gamma(\eta) = (\eta-1)! \quad \eta = 1, 2, 3, 4, \dots$$

$$\Gamma(1) = \Gamma(2) = 1 \quad \text{و} \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

تابع گاما برای مقادیر η که جدول شده است و برای سایر مقادیر $\eta > 0$ ، توسط رابطه بازگشتی $\Gamma(\eta+1) = \eta \Gamma(\eta)$ ، این تابع قابل محاسب است. میانگین و واریانس و ضریب چولگی توزیع گاما عبارتند از:

$$\Rightarrow \mu = E(X) = \frac{\eta}{\lambda}, \quad \text{var}(X) = \frac{\eta}{\lambda^2}, \quad \gamma = \frac{E(X-\mu)^3}{[\text{var}(X)]^{3/2}} = \frac{2}{\sqrt{\eta}} \quad (3)$$



توزیع گاما - ۲

این توزیع چوله به راست است و با افزایش η ضریب چوکنی آن به سمت صفر میل می‌کند، درحقیقت نمودار تابع چوکنی آن به سمت متقارن شدن نزدیک می‌شود. به ازای $\eta < 0$ ، شکل λ و به ازای $\eta > 1$ دارای یک سبک در نقطه $x = \frac{\eta-1}{\lambda}$ (مد توزیع) است. با ثابت نگه داشتن η و تغییر λ مقیاس شکل عوض می‌شود و با ثابت نگه داشتن λ و تغییر η شکل توزیع عوض می‌شود، از اینرو به λ و η به ترتیب پارامترهای مقیاس و شکل گویند

تابع جمعی توزیع گاما: (۴)
$$F_X(x) = P(X \leq x) = \int_0^x \frac{\lambda^\eta t^{\eta-1} e^{-\lambda t}}{\Gamma(\eta)} dt$$

که برای تعادری η صحیح می‌توان آن را توسط مجموع زیر بدست آورد

$$F(x) = 1 - e^{-\lambda x} \sum_{j=0}^{\eta-1} \frac{(\lambda x)^j}{j!} \quad (۵)$$

ثابت می‌شود که اگر متغیر x دارای توزیع گاما با پارامترهای η و λ باشد متغیر $y = 2\lambda x$ دارای توزیع کای-دو یا درجه آزادی $\eta = 2$ است و می‌توان توسط این خاصیت تابع جمعی توزیع گاما را با استفاده از جدول تابع جمعی توزیع x حساب کرد (درحالی که $\eta = 2$ ، صحیح نباشد باید از درونیایی خطی مربوط به دو درجه آزادی اعداد صحیح توانی طریق عدد η استفاده کرد)

توزیع گاما با پارامتر $\eta = 1$ ، توزیع نمایی با پارامتر λ است.

اگر دو متغیر مستقل x و y دارای توزیع گاما با پارامترهای شکل η و η ولی با پارامتر مشترک λ باشند، آنگاه متغیر

$$z = x + y$$

دارای توزیع گاما با پارامترهای $\eta_1 + \eta_2 = \eta$ و λ است.

استاندارد شده متغیری که دارای توزیع گاما است یعنی:

$$U = \frac{x - E(x)}{\sqrt{\text{var}(x)}}$$

برای تعادری بزرگ η دارای توزیع ممابنی $N(0,1)$ است

توزیع گاما - ۳

بر آورد پارامترهای توزیع گاما

الف - با یکاگرایی روش گشتاورها، میانگین و واریانس این توزیع به ترتیب $E(X) = \frac{\eta}{\lambda}$ و $var(X) = \frac{\eta}{\lambda^2}$ می باشد. حال اگر میانگین و واریانس آمار

استراری از جابجایی: $(x_1, x_2, \dots, x_n)$ به ترتیب $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ و $s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{1}{n-1} (\sum_{i=1}^n x_i^2 - n\bar{x}^2)$ داریم:

$$\Rightarrow \frac{\hat{\eta}}{\hat{\lambda}} = \bar{x} \quad \text{و} \quad \frac{\hat{\eta}}{\hat{\lambda}^2} = s^2 \Rightarrow \hat{\lambda} = \frac{\bar{x}}{s^2} \quad \text{و} \quad \hat{\eta} = \frac{\bar{x}^2}{s^2}$$

ب - با یکاگرایی روش حد اکثر تابع درست نمی:

$$f_x(x) = \frac{\lambda^\eta}{\Gamma(\eta)} x^{\eta-1} e^{-\lambda x}$$

$$L = \prod_{i=1}^n \frac{\lambda^\eta}{\Gamma(\eta)} x_i^{\eta-1} e^{-\lambda x_i} = \frac{\lambda^{n\eta}}{(\Gamma(\eta))^n} \left(\prod_{i=1}^n x_i \right)^{\eta-1} \cdot e^{-\lambda \sum_{i=1}^n x_i}$$

$$\log L = n\eta \log \lambda - n \log \Gamma(\eta) + (\eta-1) \log \left(\prod_{i=1}^n x_i \right) - \lambda \sum_{i=1}^n x_i$$

$$\frac{\partial \log L}{\partial \lambda} = n\eta \times \frac{1}{\lambda} - \sum_{i=1}^n x_i \quad \frac{\partial \log L}{\partial \lambda} = 0 \Rightarrow \frac{\hat{\eta}}{\hat{\lambda}} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$$

$$\Rightarrow \hat{\lambda} = \frac{\hat{\eta}}{\bar{x}} \quad (1)$$

$$\frac{\partial \log L}{\partial \eta} = n \log \lambda - n \frac{\Gamma'(\eta)}{\Gamma(\eta)} + \log \left(\prod_{i=1}^n x_i \right)$$

$$\frac{\partial \log L}{\partial \eta} = 0 \Rightarrow \log \hat{\lambda} - \frac{\Gamma'(\hat{\eta})}{\Gamma(\hat{\eta})} + \frac{1}{n} \sum_{i=1}^n \log x_i = 0 \quad (2)$$

حال برای اسامی ساده تر (۱) و (۲) را به صورت زیر می نویسیم:

$$\log \hat{\lambda} = \log \frac{\hat{\eta}}{\bar{x}} = \log \hat{\eta} - \log \bar{x} \quad (3)$$

توزیع ۶۶-۴

$\frac{1}{n} \sum_{i=1}^n \log x_i = \overline{\log x}$ همچنین براساس فرمول جیبی استرلینگ :

$$\Gamma(\eta) = \sqrt{2\pi} \eta^{\eta - \frac{1}{2}} e^{-\eta} \left(1 + \frac{1}{12\eta} + \frac{1}{288\eta^2} - \dots\right)$$

$$\approx \sqrt{2\pi} \eta^{\eta - \frac{1}{2}} e^{-\eta} \left(1 + \frac{1}{12\eta}\right)$$

$$\Rightarrow \log \Gamma(\eta) = \log \sqrt{2\pi} + \left(\eta - \frac{1}{2}\right) \log \eta - \eta + \log \left(1 + \frac{1}{12\eta}\right)$$

و با مشتق گیری از طرفین این تابع نسبت به η :

$$\frac{\Gamma'(\eta)}{\Gamma(\eta)} = \log \eta + \left(\eta - \frac{1}{2}\right) \times \frac{1}{\eta} - 1 + \frac{-\frac{1}{12\eta^2}}{1 + \frac{1}{12\eta}} \approx \log \eta - \frac{1}{2\eta} - \frac{1}{12\eta^2} \quad (4)$$

با بردن (3) و (4) در (2)

$$\log \hat{\eta} - \log \bar{x} - \left(\log \hat{\eta} - \frac{1}{2\hat{\eta}} - \frac{1}{12\hat{\eta}^2}\right) + \overline{\log x} = 0$$

$$-\log \bar{x} + \log \bar{x} + \frac{1}{2\hat{\eta}} + \frac{1}{12\hat{\eta}^2} = 0 \Rightarrow 2\eta \hat{\eta}^2 - \hat{\eta} - \frac{1}{4} = 0 \quad (5)$$

که در این معادله $y = \log \bar{x} - \log x$ و جواب نسبت معادله (5)

$$\hat{\eta} = \frac{1}{4y} \left(1 + \sqrt{1 + \frac{4y}{3}}\right) \quad (6)$$

باید توجه داشت که فرمول تقریبی استرلینگ برای مقادیر بزرگ $\hat{\eta}$ معتبر است و اگر $\hat{\eta}$ کوچک باشد این فرمول و در نتیجه فرمول (6) خالی از خطا نیست و برای مقادیر $\hat{\eta} < 5.6$ که از فرمول (6) بدست می آید جمله اصلاحی $\Delta \hat{\eta}$ براساس جدول زیر بروی فرمول (6) اعمال می شود بطوریکه :

$$\hat{\eta} = \frac{1}{4y} \left(1 + \sqrt{1 + \frac{4y}{3}}\right) - \Delta \hat{\eta} \quad (7)$$

دیس :

$$\hat{\lambda} = \frac{\hat{\eta}}{\bar{x}}$$

$$\bar{y} = \log \bar{x} - \log x \quad (8)$$

توزیع ۶۶-۵

$\hat{\eta}$	$\Delta \hat{\eta}$	η	$\Delta \hat{\eta}$	$\hat{\eta}$	$\Delta \hat{\eta}$
۰.۲	۰/۰۳۴	۱,۰	۰/۰۰۹	۱,۸	۰/۰۰۴
۰.۳	۰/۰۲۹	۱,۱	۰/۰۰۸	۱,۹	۰/۰۰۳
۰.۴	۰/۰۲۵	۱,۲	۰/۰۰۷	۲,۲	۰/۰۰۳
۰.۵	۰/۰۲۱	۱,۳	۰/۰۰۶	۲,۳	۰/۰۰۲
۰.۶	۰/۰۱۷	۱,۴	۰/۰۰۶	۳,۱	۰/۰۰۲
۰.۷	۰/۰۱۴	۱,۵	۰/۰۰۵	۳,۲	۰/۰۰۱
۰.۸	۰/۰۱۲	۱,۶	۰/۰۰۵	۵,۵	۰/۰۰۱
۰.۹	۰/۰۱۱	۱,۷	۰/۰۰۴	۵,۶	۰/۰۰۰

جدول جبهه اصلاحی برآورد پارامتر η

فرمول کمی زیر توسط گرین و دُراند (Greenwood and Durand) برای برآورد پارامتر کمی η در بدست آمده است.

$$(a) \begin{cases} \hat{\eta} = \frac{1}{y} (0.5000876 + 0.1648852y - 0.544274y^2) : 0.5 \leq y \leq 1 \\ \hat{\eta} = \frac{1.898919 + 9.05995y + 0.9775373y^2}{y(17.79728 + 11.968477y + y^2)} : 0.5772 \leq y \leq 17 \end{cases} \Rightarrow \hat{\lambda} = \frac{\hat{\eta}}{x}$$

برآورد پارامترهای توزیع کاما توسط روش حداکثر تابع درست نمایی، ناپایب
ثبیت و برای آن کمی

$$E(\hat{\eta} - \eta) = \frac{3\eta}{n}$$

$$E(\hat{\eta} - \eta) \approx \frac{3\hat{\eta}}{n}$$

و با تقریب
که می توان به $\hat{\eta}$ بدست آمده توسط فرمول کمی (۷) یا (۹) با افزودن
جبهه $\frac{3\hat{\eta}}{n}$ ، آن را اصلاح کرد و پس توسط $\hat{\lambda} = \frac{\hat{\eta}}{x}$ ، برآورد
در بدست آورد.

توزیع گاما - ۶-۴
 مثال - میزان بارندگی سالانه در یک منطقه از توزیع گاما پیروی می کند. برای مدت ۱۸ سال آمار زیر درست داریم. اعداد بر حسب اینج هستند.

۱۳,۲۶	۷,۷۰	۱۵,۱۷
۳,۱۳	۱۷,۶۴	۱۰,۴۰
۱۵,۱۷	۲۲,۹۱	۱۸,۰۲
۱۵,۵۰	۱۸,۸۹	۱۶,۲۵
۱۴,۲۲	۱۲,۸۲	۱۱,۷۷
۲۱,۲۰	۱۱,۳۸	۱۷,۹۳

پارامترهای این توزیع را با روش گستردها و حداکثر تابع درست نمایی برآورد کنند.
 احتمال اینکه میزان بارندگی یک سال از ۲۰ این تجاوز نکند بدست آورده

$$\text{حل - } \sum_{i=1}^{18} x_i = ۲۹۳,۷۳ \text{ in} \quad , \quad \sum_{i=1}^{18} x_i^2 = ۴۲۴۹,۱۵ \text{ in}$$

الف - برآورد پارامتریهای λ و η در توزیع با تابع چگالی احتمال:
 $f(x) = \frac{\lambda^\eta}{\Gamma(\eta)} x^{\eta-1} e^{-\lambda x} \quad x > 0$
 $(\lambda, \eta > 0)$ با روش گستردها:

$$\Rightarrow \bar{x} = \frac{\sum x_i}{n} = \frac{۲۹۳,۷۳}{۱۸} = ۱۶,۳۲ \text{ in}$$

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{1}{n-1} (\sum x_i^2 - n\bar{x}^2)}$$

$$= \sqrt{\frac{1}{17} (۴۲۴۹,۱۵ - \frac{1}{18} (۲۹۳,۷۳)^2)} = ۴,۷۵ \text{ in}$$

$$\Rightarrow \hat{\lambda} = \frac{\bar{x}}{s^2} = \frac{۱۶,۳۲}{(۴,۷۵)^2} = ۰,۷۴۹ \text{ (in)}^{-1}$$

$$\hat{\eta} = \frac{\bar{x}^2}{s^2} = \frac{۲۶۱,۹۲}{۲۲,۵۶} = ۱۱,۶۱۳$$

ب - برآورد پارامترها با روش حداکثر تابع درست نمایی:

$$\ln x = \left(\sum_{i=1}^n \ln x_i \right) / n = \frac{\sum_{i=1}^{18} \ln x_i}{18} = ۲,۹۱۲, \quad \ln \bar{x} = \ln ۱۶,۳۲ = ۲,۹۱۳$$

توزیع کابا ۶-۷

$$y = \ln \bar{x} - \ln x = 2,914 - 2,912 = 0,002$$

$$\hat{\eta} = \frac{1}{4y} \left(1 + \sqrt{1 + \frac{4y}{3}} \right) = \frac{1}{4(0,002)} \left(1 + \sqrt{1 + 1,333(0,002)} \right) = 7,107$$

چون $\hat{\eta} = 7,107 > 5,6$ احتیاج به اصلاح ندارد

$$\hat{\lambda} = \frac{\hat{\eta}}{\bar{x}} = \frac{7,107}{14,95} = 0,475$$

و بر اساس روش گریس وود و درانز با توجه به اینکه $0,002 < 0,05$ برآوردی $\hat{\eta} = 0,002$ بدست می آوریم

$$\hat{\eta} = \frac{0,000876 + 0,1648852y - 0,54427y^2}{y} = 7,107$$

$$\hat{\lambda} = \frac{\hat{\eta}}{\bar{x}} = \frac{7,107}{14,95} = 0,475$$

برآورد $\hat{\eta}$ توسط روش حداکثر تابع درست نشی دارای بزرگی $E(\hat{\eta} - \eta) = \frac{3\eta}{n} \approx \frac{3\hat{\eta}}{n}$ یعنی $\frac{3 \times 7,107}{n} = 1,185$ است. پس برآورد اصلاح شده برای ایجاد ناریبی:

$$\hat{\eta} = 7,107 - 1,185 = 5,922$$

$$\hat{\lambda} = \frac{\hat{\eta}}{\bar{x}} = \frac{5,922}{14,95} = 0,396$$

حساب احتمال کسبه میزان بارندگی سالانه از ۲۰ in تجاوز کند:

$$P(X > 20,00) = 1 - P(X \leq 20,00) = 1 - \int_0^{20} \frac{\lambda^2 x^{2-1} e^{-\lambda x}}{\Gamma(2)} dx$$

$$= 1 - \left(\text{سطح زیر منحنی احتمالی توزیع } x \text{ بارندگی آزادگی } 1,185 \text{ را } x = 20 \text{ بده} \right) = 1 - 0,824 = 0,176$$

$$\text{د طول: } x^2 = 2\lambda x = 2 \times 0,396 \times 20 = 15,84$$

توزیع مقدار اکسترم (Extreme value Distributions) - ۱

۱- طرح مسأله: جامعه آماری P (جامعه مادر) با توزیع پیوسته با تابع چگالی احتمال

$f_X(x)$ و تابع تجمعی $F_X(x) = P(X \leq x)$ ، مفروض است. از این جامعه یک

نمونه گیری به حجم n : $\{X_1, X_2, \dots, X_n\}$ صورت می گیرد. توزیع متغیرهای

تصادفی $Y = \max\{X_1, X_2, \dots, X_n\}$ و $Z = \min\{X_1, X_2, \dots, X_n\}$ را توزیع

تصادفی اکسترم جامعه P می نامیم. میخواهیم توزیع متغیرهای Z و Y را تعیین کنیم

برای هر دو توزیع ابتدا تابع تجمعی این توزیع ها را بداییم کنیم.

الف- توزیع $Y = \max\{X_1, X_2, \dots, X_n\}$:

$$F_Y(y) = P(Y \leq y) = P(\max\{X_1, X_2, \dots, X_n\} \leq y)$$

$$= P(X_1 \leq y, X_2 \leq y, \dots, X_n \leq y)$$

$$= P(X_1 \leq y) P(X_2 \leq y) \times \dots \times P(X_n \leq y) = (F_X(y))^n$$

زیرا X_i مستقل از هم هستند و این احتمالات باهم برابرند و هر کدام برابر

$$F_X(y) \text{ است. لذا } F_Y(y) = (F_X(y))^n$$

و تابع چگالی احتمال متغیر Y :

$$f_Y(y) = F_Y'(y) = n(F_X(y))^{n-1} F_X'(y) = n(F_X(y))^{n-1} f_X(y)$$

مثال ۱- فرض کنید مدت زمان بین دو بارندگی دارای توزیع نمایی با میانگین ۴ روز است و فرض کنید زمان بین دو بارندگی مستقل از سایر زمان های مشابه دیگر است (۱). توزیع زمان تاگسیم فاصله بین دو بارندگی در ده بارندگی را بدست آورید. (۲). احتمال اینکه از یک زمان از هشت روز بجا نماند، حساب کنید. حل (۱) تابع چگالی توزیع نمایی عبارت از $f_X(x) = \lambda e^{-\lambda x}$ و $\lambda > 0$

میانگین این توزیع $\mu = \frac{1}{\lambda}$ است. پس $\lambda = \frac{1}{4}$ و $f_X(x) = \frac{1}{4} e^{-\frac{1}{4}x}$ $x > 0$. $F_X(x) = \int_0^x \frac{1}{4} e^{-\frac{1}{4}x} dx = 1 - e^{-\frac{x}{4}}$

$$f_Y(y) = 9(F_X(y))^8 f_X(y)$$

$$Y = \max\{X_1, X_2, \dots, X_9\}$$

توزیع تعدادی اکسپوننسیل

$$f_Y(y) = \frac{\lambda}{2} e^{-\lambda/2} (1 - e^{-\lambda/2})^{\lambda}$$

یعنی $f_Y(y) = \frac{\lambda}{2} (1 - e^{-\lambda/2})^{\lambda} \times \frac{1}{2} e^{-\lambda/2}$

(۲)

$$P(Y > 1) = 1 - P(Y \leq 1) = 1 - F_Y(1) = 1 - (F_X(1))^9$$

$$= 1 - (1 - e^{-\frac{1}{2}})^9 = 1 - (1 - e^{-0.5})^9 = 1 - 0.170168 \approx 0.829832$$

ب - توزیع $Z = \min\{X_1, X_2, \dots, X_n\}$

$$F_Z(z) = P(\min\{X_1, X_2, \dots, X_n\} \leq z) = 1 - P(\min\{X_1, X_2, \dots, X_n\} \geq z)$$

$$= 1 - P(X_1 \geq z \text{ و } X_2 \geq z \text{ و } \dots \text{ و } X_n \geq z) = 1 - P(X_1 \geq z)P(X_2 \geq z) \dots P(X_n \geq z)$$

$$= 1 - [P(X \geq z)]^n = 1 - [1 - P(X \leq z)]^n = 1 - (1 - F_X(z))^n$$

و تابع چگالی احتمال میسر z :

$$f_Z(z) = n(1 - F_X(z))^{n-1} F'_X(z)$$

و یا $f_Z(z) = n(1 - F_X(z))^{n-1} f_X(z)$

به عنوان مثال وقتی جامعه دارای توزیع نمایی با پارامتر λ باشد:

داریم $f_X(x) = \lambda e^{-\lambda x}$ و $F_X(x) = 1 - e^{-\lambda x}$ و تابع چگالی احتمال

توزیع $Z = \min(X_1, X_2, \dots, X_n)$ عبارت از:

$$f_Z(z) = n(e^{-\lambda z})^{n-1} \times \lambda e^{-\lambda z} = n\lambda e^{-n\lambda z}$$

دیده می شود که Z هم دارای توزیع نمایی با پارامتر $n\lambda$ است.

۲ - توزیع جهانی تعدادی اکسپوننسیل برای تعدادی بزرگ n .

بر حسب توزیع جامعه ای که از آن نمونه گیری می شود (جامعه مادر) و نوع اکسپوننسیل (ماکسیم و مینیم) این توزیع (یعنی توزیع تعدادی اکسپوننسیل) به سه نوع طبقه بندی می شود.

توزیع تعدادی اکسیریم - ۳

نوع I - توزیع جامعه مادر درجهت نوع اکسیریم نامحدود است و ناممکن است درازی
توزیع وجود دارند (توزیع‌های نوع نمایی)

۱ - نوع اکسیریم ماکسیم ، توزیع‌های جامعه مادر : نرمال ، لگرنرمال ، نمایی ، دگاما

۲ - نوع اکسیریم می‌شیم ، توزیع جامعه مادر : نرمال

نوع II - توزیع جامعه مادر درجهت نوع اکسیریم نامحدود است ، ولی تمام گستره‌های
توزیع وجود ندارند (توزیع‌های نوع کوئی)

۱ - نوع اکسیریم ماکسیم و می‌شیم و توزیع‌های جامعه مادر توزیع کوئی - و توزیع t - استیودنت

نوع III - توزیع جامعه مادر درجهت نوع اکسیریم محدود است .

۱ - نوع اکسیریم ماکسیم و توزیع جامعه مادر : توزیع بتا

۲ - نوع اکسیریم می‌شیم و توزیع‌های جامعه مادر : لگرنرمال ، گاما و نمایی

توزیع تعدادی اکسیریم نوع II به ندرت در پدیده‌های هیدرولوژیکی بکار می‌رود

کاربردی که این نوع توزیع تعدادی اکسیریم ، نوع I برای تعداد ماکسیم
و نوع III برای تعداد می‌شیم است ، چون کمیت‌های تعدادی در

هیدرولوژی دوست راست نامحدود به لحاظ است بودن آنها
در سمت چپ محدود و دارای کران پایینی صفر می‌باشند

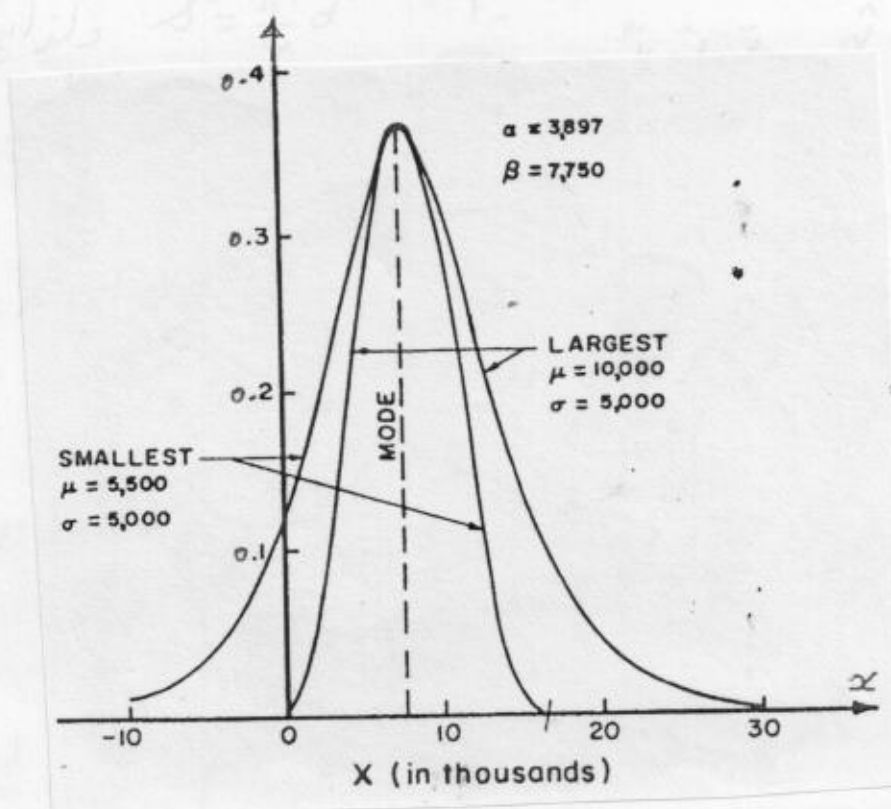
۳ - توزیع گمبل (Gumbel's distribution)

توزیع مجانبی نوع I برای تعداد ماکسیم و می‌شیم دارای تابع چگالی احتمال

$$f_x(x) = \frac{1}{\alpha} e^{-\frac{(x-\beta)}{\alpha}} - e^{-\frac{(x-\beta)}{\alpha}} \quad \begin{matrix} -\infty < x < \infty \\ \alpha > 0 \text{ و } -\infty < \beta < \infty \end{matrix}$$

به طوریکه علامت - (علامت بالا) در تابع چگالی احتمال تعداد ماکسیم
و علامت + (علامت پایین) برای تابع چگالی احتمال تعداد می‌شیم بکار
میرود.

توزیع تعادری اکسپوننسیال - ۴
 α و β به ترتیب پارامترهای مقیاس و مکان می باشند، برای هر دو
 توزیع، β مد توزیع می باشد. نمودارهای چگالی احتمال توزیع های تعادری
 یکسوم و دو سوم نسبت به خط $\alpha = \beta$ قرینه یکدیگر می باشند. در زیر آن ها
 به ازای $\alpha = 3897$ و $\beta = 7750$ رسم شده اند.



میانگین و واریانس توزیع های تعادری اکسپوننسیال نوع I عبارتند از:

$$\begin{aligned} \mu_x = E(X) &= \beta + 0.577\alpha && \text{برای تعادری یکسوم} \\ &= \beta - 0.577\alpha && \text{برای تعادری دو سوم} \end{aligned}$$

$$\text{Var}(X) = \frac{\pi^2}{6} \alpha^2 \approx 1.645 \alpha^2$$

عدد ۰.۵۷۷ در بالا صرفاً ثابت اویلر-ماسکرونی است و عبارت از

$$\gamma = \lim_{n \rightarrow \infty} \left(\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} - \ln n \right) = -\pi'(1) \approx 0.577$$

توزیع مقادیر اکسیریم ۶-
 پس از محاسبه $\hat{\alpha}$ توسط رابطه (۱۷)، $\hat{\beta}$ بدست می آید.
 تابع چگلی توزیع گمبل

$$F_X(x) = P(X \leq x) = \frac{1}{\alpha} \int_{-\infty}^x e^{-\frac{(x-\beta)}{\alpha}} - e^{-\frac{(x-\beta)}{\alpha}} dx$$

$$= \frac{1}{\alpha} \int_{-\infty}^x e^{-\frac{(x-\beta)}{\alpha}} \times e^{-\frac{(x-\beta)}{\alpha}} dx$$

$$= \begin{cases} e^{-e^{-\frac{(x-\beta)}{\alpha}}} & (\text{maximum}) \\ 1 - e^{-e^{-\frac{(x-\beta)}{\alpha}}} & (\text{minimum}) \end{cases}$$

سوال ۲- اکسیم ری سبل سالانه از توزیع گمبل (ماکسیم) تبعیت می کند.
 مطلوبیت محاسبه دوره بازگشت سبل ماکسیم سالانه که بزرگتر از میانگین
 توزیع باشد

$$P(X \leq \mu) = \int_{-\infty}^{\mu = \beta + 0.577\alpha} \frac{1}{\alpha} e^{-\frac{x-\beta}{\alpha}} \times e^{-\frac{x-\beta}{\alpha}} dx = e^{-e^{-\frac{x-\beta}{\alpha}}} \Big|_{-\infty}^{\beta + 0.577\alpha}$$

$$= e^{-e^{-\frac{0.577\alpha}{\alpha}}} = e^{-e^{-0.577}} = 0.570321 \Rightarrow P(X \geq \mu) = 0.4296$$

$$\text{سال} = T = \frac{1}{0.4296} = 2.327 \approx 2.33$$

سوال ۳- ثابت کنید که برای توزیع مقادیر ماکسیم $\mu = \beta + 0.577\alpha$

حل:

$$\mu = E(X) = \int_{-\infty}^{\infty} x \times \frac{1}{\alpha} e^{-\frac{(x-\beta)}{\alpha}} \times e^{-\frac{(x-\beta)}{\alpha}} dx$$

برای محاسبه این انگرال تغییر متغیر $e^{-\frac{x-\beta}{\alpha}} = y$ بکار می ریم.

توزیع مقادیر اکسپوننسیال - ۷

$$\Rightarrow -\frac{1}{\alpha} e^{-\frac{x-\beta}{\alpha}} dx = dy \Rightarrow \frac{1}{\alpha} e^{-\frac{x-\beta}{\alpha}} dx = -dy$$

$$-\frac{x-\beta}{\alpha} = \ln y \Rightarrow x = \beta - \alpha \ln y$$

$$x \rightarrow -\infty \Rightarrow y \rightarrow e^{\infty} \Rightarrow y \rightarrow +\infty ; x \rightarrow +\infty \Rightarrow y \rightarrow e^{-\infty} = 0$$

$$\mu = \int_{-\infty}^{\infty} (\beta - \alpha \ln y) e^{-y} (-dy) = \int_0^{\infty} (\beta - \alpha \ln y) e^{-y} dy$$

$$\Gamma'(p) = \int_0^{\infty} y^{p-1} e^{-y} dy \quad \Leftarrow \quad \int_0^{\infty} \beta e^{-y} dy = \beta$$

$$\Rightarrow \Gamma'(p) = \int_0^{\infty} y^{p-1} \ln y e^{-y} dy \Rightarrow \Gamma'(1) = \int_0^{\infty} \ln y e^{-y} dy$$

$$\text{یعنی} \quad \int_0^{\infty} \ln y e^{-y} dy = \Gamma'(1) = -\gamma$$

$$\mu = \beta - \alpha \int_0^{\infty} \ln y \cdot e^{-y} dy = \beta - \alpha(-\gamma) = \beta + \alpha\gamma = \beta + 0.577\alpha$$

۴ - توزیع اکسپوننسیال - توزیع وایبول (Weibull's distribution)

توزیع مقادیر میثیم وقتی که توزیع جامعه مادر در طرف چپ محدود (دارای کران پایینی) و حجم نمونه (n) بزرگ باشد، دارای توزیعی است با نام توزیع وایبول (دوباراکس) یا تابع چگالی احتمال

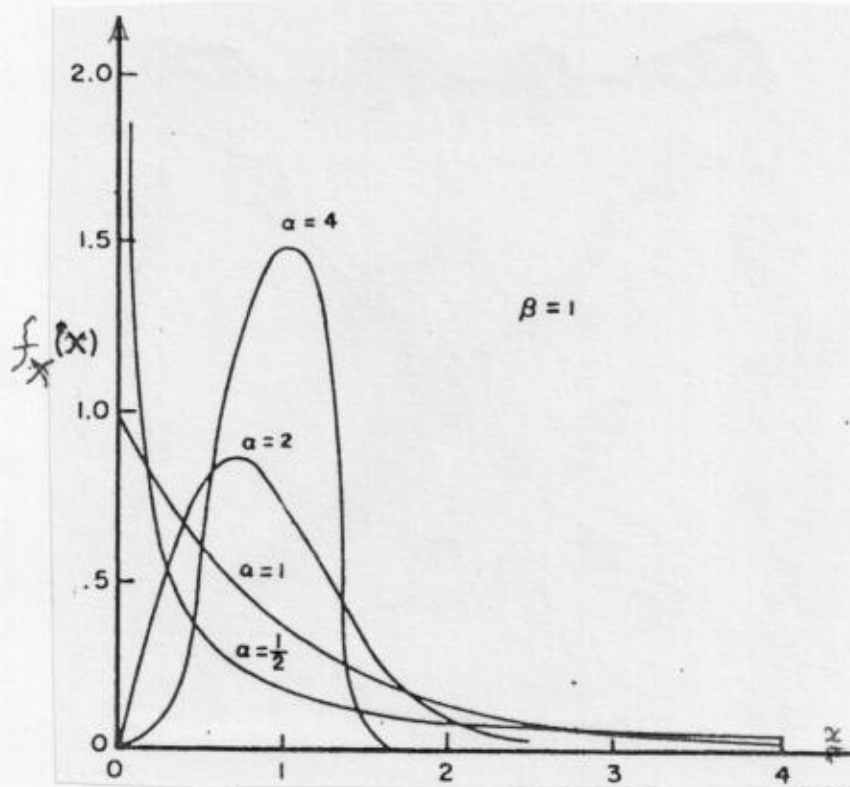
$$f_X(x) = \alpha x^{\alpha-1} \beta^{-\alpha} e^{-\left(\frac{x}{\beta}\right)^{\alpha}} \quad x \geq 0 \text{ و } \alpha, \beta > 0$$

$$F_X(x) = P(X \leq x) = \int_0^x \alpha x^{\alpha-1} \cdot \frac{1}{\beta^{\alpha}} e^{-\left(\frac{x}{\beta}\right)^{\alpha}} dx = -e^{-\left(\frac{x}{\beta}\right)^{\alpha}} \Big|_0^x = 1 - e^{-\left(\frac{x}{\beta}\right)^{\alpha}}$$

بایگین، داریانس، و غیره چوگلی این توزیع را می توان با یک اشتدالی یعنی $E(X^r)$ را حساب کرد

$$E(X^r) = \int_0^{\infty} x^r f_X(x) dx = \int_0^{\infty} \alpha x^{r+\alpha-1} \beta^{-\alpha} e^{-\left(\frac{x}{\beta}\right)^{\alpha}} dx \quad \left(\frac{x}{\beta}\right)^{\alpha} = y \Rightarrow x = \beta y^{\frac{1}{\alpha}}$$

توزیع معادیر اکسیرم - ۱



$$\Rightarrow E(X^r) = \int_0^{\infty} \alpha (\beta y^{\frac{1}{\alpha}})^{r+\alpha-1} \beta \cdot e^{-y} \frac{\beta}{\alpha} y^{\frac{1}{\alpha}-1} dy$$

$$= \beta^{r+\alpha-1} \cdot \int_0^{\infty} y^{\frac{r+\alpha}{\alpha}-1} e^{-y} dy = \beta^r \Gamma(1 + \frac{r}{\alpha})$$

نکته: $dx = \frac{\beta}{\alpha} y^{\frac{1}{\alpha}-1} dy$

$$\mu = E(X) = \beta \Gamma(1 + \frac{1}{\alpha})$$

$$Var(X) = E(X^2) - \mu^2 = \beta^2 \Gamma(1 + \frac{2}{\alpha}) - \beta^2 \Gamma^2(1 + \frac{1}{\alpha})$$

$$= \beta^2 [\Gamma(1 + \frac{2}{\alpha}) - \Gamma^2(1 + \frac{1}{\alpha})]$$

$$\gamma = \frac{E(X-\mu)^3}{[Var(X)]^{\frac{3}{2}}} = \frac{E(X^3 - 3\mu X^2 + 3\mu^2 X - \mu^3)}{[Var(X)]^{\frac{3}{2}}} = \frac{E(X^3) - 3\mu E(X^2) + 3\mu^2 E(X) - \mu^3}{[Var(X)]^{\frac{3}{2}}}$$

$$= \frac{E(X^3) - 3\mu E(X^2) + 3\mu^2}{[Var(X)]^{\frac{3}{2}}} = \frac{\Gamma(1 + \frac{3}{\alpha}) - 3\Gamma(1 + \frac{1}{\alpha})\Gamma(1 + \frac{2}{\alpha}) + 3\Gamma^3(1 + \frac{1}{\alpha})}{[\Gamma(1 + \frac{2}{\alpha}) - \Gamma^2(1 + \frac{1}{\alpha})]^{\frac{3}{2}}}$$

در این کسر عدد ۳ از صورت و خارج آن حذف می شود.

بر آورد پارامترهای α و β از محل دستگاه معادلات زیر بدست می آید:

توزیع مقادیر اکسپوننسیال

$$\begin{cases} \bar{X} = \hat{\beta} \Gamma(1 + \frac{1}{\alpha}) \\ s^2 = \hat{\beta}^2 \left(\Gamma(1 + \frac{2}{\alpha}) - \Gamma^2(1 + \frac{1}{\alpha}) \right) \end{cases}$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$$

$$\Rightarrow \frac{s^2}{\bar{X}^2} = \frac{\Gamma(1 + \frac{2}{\alpha}) - \Gamma^2(1 + \frac{1}{\alpha})}{\Gamma^2(1 + \frac{1}{\alpha})} \Rightarrow \frac{\Gamma(1 + \frac{2}{\alpha})}{\Gamma^2(1 + \frac{1}{\alpha})} = \frac{s^2}{\bar{X}^2} + 1$$

با حل این معادله به صورت سعی و خطا $\hat{\alpha}$ حساب می شود و سپس $\hat{\beta} = \frac{\bar{X}}{\Gamma(1 + \frac{1}{\alpha})}$ برای برآورد پارامترها با جایگزینی روش حداکثر تابع درست نمایی: ابتداء می گیریم $\eta = \beta^{-\alpha}$

$$f_X(x) = \alpha x^{\alpha-1} \eta e^{-\eta x^\alpha}$$

$$\ln L = n \ln \alpha + (\alpha-1) \sum_{i=1}^n \ln X_i + n \ln \eta - \eta \sum_{i=1}^n X_i^\alpha$$

$$\frac{\partial (\ln L)}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^n \ln X_i - \eta \sum_{i=1}^n X_i^\alpha \ln X_i = 0$$

$$\frac{\partial (\ln L)}{\partial \eta} = \frac{n}{\eta} - \sum_{i=1}^n X_i^\alpha = 0$$

از معادله اخیر نتیجه می شود:

$$\hat{\eta} = \frac{n}{\sum_{i=1}^n X_i^\alpha} \quad \text{و با بردن } \hat{\eta} \text{ در معادله اول}$$

وساده نمودن آن، این معادله روی $\hat{\alpha}$ به دست می آید که خیلی پیچیده است

$$\hat{\alpha} = \frac{1}{\frac{\sum_{i=1}^n X_i^\alpha \ln X_i}{\sum_{i=1}^n X_i^\alpha} - \frac{1}{n} \sum_{i=1}^n \ln X_i}$$

که بارش سعی و خطا حل می شود، پس $\hat{\eta}$ و در نهایت توسط رابطه $\hat{\alpha} = \hat{\eta}^{-\frac{1}{\beta}}$ ، $\hat{\beta}$ محاسبه می شود.

به ازای $\alpha=1$ ، توزیع وایبول به توزیع نمایی منجر می شود و با افزایش α نمودار توزیع به سمت متعادل شدن میل می کند.

توزیع مقدار آکسیرم - ۱۰
اگر کران پائین جامعه بجای صفر مقدار $\epsilon > 0$ باشد. توزیع وایبول
به پارامتری با تابع چگالی احتمال:

$$f_X(x) = \alpha(x-\epsilon)^{\alpha-1}(\beta-\epsilon)^{-\alpha} \exp\left\{-\left(\frac{x-\epsilon}{\beta-\epsilon}\right)^\alpha\right\} \quad x > \epsilon, \alpha > 0, \beta > \epsilon$$

تابع تجزیه، میانگین، واریانس این توزیع عبارتند از:

$$F_X(x) = P(X \leq x) = 1 - e^{-\left(\frac{x-\epsilon}{\beta-\epsilon}\right)^\alpha}$$

$$\mu = \epsilon + (\beta - \epsilon) \Gamma\left(1 + \frac{1}{\alpha}\right)$$

$$\sigma^2 = \text{Var}(X) = (\beta - \epsilon)^2 \left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \Gamma^2\left(1 + \frac{1}{\alpha}\right) \right]$$

ضریب چوگلی در توزیع وایبول به پارامتری همان ضریب چوگلی
برای $\epsilon = 0$ است و فقط بر حسب α بیان می شود.
از دو رابط بالا نتیجه می شود:

$$\begin{cases} \beta = \mu + \sigma A(\alpha) \\ \epsilon = \beta - \sigma B(\alpha) \end{cases}$$

که $A(\alpha) = \left[1 - \Gamma\left(1 + \frac{1}{\alpha}\right)\right]^{-\frac{1}{\alpha}}$ و $B(\alpha) = \left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \Gamma^2\left(1 + \frac{1}{\alpha}\right)\right]^{-\frac{1}{\alpha}}$

حال اگر بخواهیم با یکاگیری روش گشتاورها برای یک نمونه $\{X_1, X_2, \dots, X_n\}$ از جامعه X و $\bar{X}$ و s^2 و s^4 به ترتیب برآوردهای
برای μ و σ^2 و σ^4 (ضریب چوگلی) باشند: بطوریکه

$$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad \hat{\sigma}^2 = s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}, \quad \hat{\sigma}^4 = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^4}{s^4}$$

ابتداءً با حل معادله

$$\hat{\sigma}^4 = \frac{\Gamma\left(1 + \frac{4}{\alpha}\right) - 3\Gamma\left(1 + \frac{2}{\alpha}\right)\Gamma\left(1 + \frac{1}{\alpha}\right) + 2\Gamma^3\left(1 + \frac{1}{\alpha}\right)}{\left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \Gamma^2\left(1 + \frac{1}{\alpha}\right)\right]^2}$$

روی α ، پارامتر α را یک عدد می گیریم. توسط جدول صفر بعد α و عبارتهای
 $A(\alpha)$ و $B(\alpha)$ تعیین می شود (در صورت لازم از روشهای مختلفی استفاده
می کنیم). دست آخر از روابط $\hat{\mu} = \epsilon + \sigma A(\alpha)$ و $\hat{\sigma}^2 = \sigma^2 B(\alpha)$ و $\epsilon = \hat{\mu} - \hat{\sigma} A(\alpha)$ (تویا راسخ

توزیع تعداد اکسیریم - 11

γ	$1/\alpha$	$A(\alpha)$	$B(\alpha)$
-1.000	0.02	0.446	40.005
-0.971	0.03	0.444	26.987
-0.917	0.04	0.442	20.481
-0.867	0.05	0.439	16.576
-0.638	0.10	0.425	8.737
-0.254	0.20	0.389	4.755
0.069	0.30	0.346	3.370
0.359	0.40	0.297	2.634
0.631	0.50	0.246	2.159
0.896	0.60	0.193	1.815
1.160	0.70	0.142	1.549
1.430	0.80	0.092	1.334
1.708	0.90	0.044	1.154
2.000	1.00	0.000	1.000
2.309	1.10	-0.040	0.867
2.640	1.20	-0.077	0.752
2.996	1.30	-0.109	0.652
3.382	1.40	-0.136	0.563
3.802	1.50	-0.160	0.486
4.262	1.60	-0.180	0.418
4.767	1.70	-0.196	0.359
5.323	1.80	-0.208	0.308
5.938	1.90	-0.217	0.263
6.619	2.00	-0.224	0.224
7.374	2.10	-0.227	0.190
8.214	2.20	-0.229	0.161

دیگر برآورد می شود.
سال ۴ - می نیم دبی سالانه یک آبراه دارای توزیع وایبول بسیار تنگی

می باشد. براساس یک نمونه گیری $\bar{x} = 125$ cfs و $s = 50$ cfs و ضریب
چوگس ۱٫۴ برآورد شده است. احتمال اینکه می نیم دبی سالانه کمتر از
۱۰۰ cfs باشد، چقدر است.

حل - از روی جدول بالا و یک درونیابی بدست می آوریم :

$$\Rightarrow \hat{\alpha} = 1.4 \Rightarrow \frac{1}{\hat{\alpha}} = 0.714, \hat{\alpha} = 1.296, A(\hat{\alpha}) = 0.098, B(\hat{\alpha}) = 1.296$$

$$\hat{\beta} = \hat{\mu} + \hat{\sigma} A(\hat{\alpha}) = 125 + 50(0.098) = 129.9$$

$$\hat{\sigma} = \hat{\mu} - \hat{\sigma} B(\hat{\alpha}) = 125 - 50(1.296) = 61.9$$

$$P(X < 100) = F_X(100) = 1 - \exp\left\{-\left(\frac{\alpha - \epsilon}{\beta - \epsilon}\right)^\alpha\right\} = 1 - e^{-\left(\frac{100 - 61.9}{129.9 - 61.9}\right)^{1.296}} = 1 - e^{-0.48} = 0.381$$

$$P(X < 100) = 0.381$$

حدود یا بازه اعتماد (Confidence Interval) - ۱

- ۱- مقدمه - اگر یک معادله جبری (یا غیر جبری) خیالی پیچیده باشد که نتوان در آن روش‌های دقیق حساب کرد، توسط یکی از روشهای عددی (مثلاً روش رافسون - نیوتون) مقدار تقریبی آن را تا هر دقتی که خواسته باشیم، می‌شود حساب کرد. به عنوان مثال اگر برای معادله‌ای، روش تقریبی ۲,۷۵۹ تا سه رقم اعشار (گرد شده) بدست آمده باشد، معنی آن این است که روش دقیق معادله در بازه $[۲,۷۵۷, ۲,۷۵۵]$ قرار دارد. بنابراین می‌توانیم چنین سانه‌ای را وقتی که پارامتر مفروض α از توزیع براساس نمونه استخرهای $X_1, X_2, \dots, X_n$ از جامعه برآورد کرده ایم حل کنیم. معنی بازه اعتماد $[L, U]$ تعیین کنیم که تعداد واقعی پارامتر α درون آن قرار گیرد. از آنجایی که محاسبه کرانه‌های این بازه معنی L و U هم مانند α از روی آمار استخرهای از جامعه برآورد می‌شوند متغیرهای تصادفی هستند و ممکن است بازه $[L, U]$ به طور یقین مقدار واقعی α را دربرگیرد و با اطمینان این امر سرگردانند مثلاً ۹۵٪ یا ۹۹٪. که گفته می‌شود $[L, U]$ بازه اعتماد ۹۵٪ (یا ۹۹٪) پارامتر α است. به طور دقیق‌تر تعریف بازه اعتماد $(1-\alpha)$ درصد (۴) تعداد است (تست و کوئیک) پارامتر α عبارتست از بازه $[L, U]$ با کوچکترین طول $(U-L)$ به طوری که با احتمال $1-\alpha$ پارامتر α درون این بازه قرار گیرد. معنی $P(L \leq \alpha \leq U) = 1-\alpha$
- ۲- محاسبه حدود اعتماد معنی محاسبه L و U برای احتمال مفروض $1-\alpha$ قبل از قبول بندی در حالت کلی، این سانه را برای تعیین بازه اعتماد پارامتر μ از توزیع $N(\mu, \sigma)$ وقتی که σ معلوم باشد (کوئیک) یا برآورد آن نیست (۱) و μ را برآورد کرده باشیم حل می‌کنیم.

بازده اعتماد ۲-

بر اساس نمونه $\{X_1, X_2, \dots, X_n\}$ از جامعه $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

برآورد μ است و می دانیم که $\hat{\mu} = \bar{X}$ متغیر است تصادفی که دارای

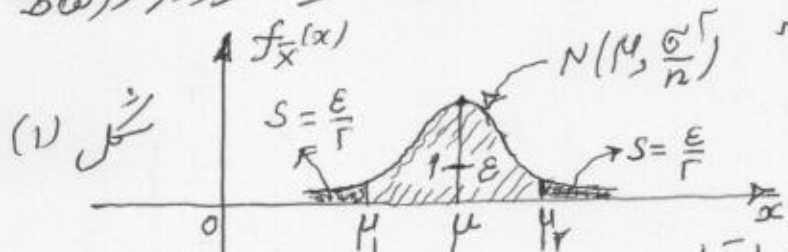
توزیع $N(\mu, \frac{\sigma^2}{n})$ است. ابتدا دو عدد μ_1 و μ_2 را چنان تعیین

می کنیم بطوریکه $P(\mu_1 \leq \hat{\mu} \leq \mu_2) = 1 - \epsilon$ و $\mu_1 - \mu_2$ کمترین

تعداد ممکن باشد. بر اساس لیمای که در زیر ثابت خواهیم کرد مقدار

$\mu_1 - \mu_2$ ، وقتی که سطح زیر نمودار تابع چگالی احتمال $N(\mu, \frac{\sigma^2}{n})$ بین طول های μ_1 و μ_2 مقدار مشخص $1 - \epsilon$ است، می بینیم است که عرض های این نمودار در نقاط μ_1 و μ_2 با هم برابر باشد.

و چون این توزیع متقارن است معادل است



شکل (۱)

که سطح زیر منحنی تابع چگالی احتمال متغیر $\bar{X}$ از $-\infty$ تا μ_1 برابر است با مساحت زیر منحنی از μ_2 تا $+\infty$ که برابرند با $\frac{\epsilon}{2}$ (حتی $\mu_1 + \mu_2 = 2\mu$)

برمی گردیم به محاسبه μ_1 و μ_2 :

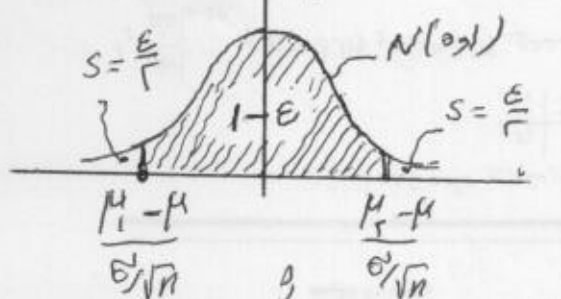
$$P(\mu_1 \leq \hat{\mu} \leq \mu_2) = P\left(\frac{\mu_1 - \mu}{\sigma/\sqrt{n}} \leq \frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} \leq \frac{\mu_2 - \mu}{\sigma/\sqrt{n}}\right) = 1 - \epsilon$$

ولی

$$\frac{\hat{\mu} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1) \text{ و } \frac{\mu_1 - \mu}{\sigma/\sqrt{n}} = -\frac{\mu_2 - \mu}{\sigma/\sqrt{n}}$$

$$(\mu - \mu_1 = \mu_2 - \mu)$$

و طبق شکل (۲):



شکل (۲)

$$\frac{\mu_1 - \mu}{\sigma/\sqrt{n}} = -z_{1-\frac{\epsilon}{2}} \quad \frac{\mu_2 - \mu}{\sigma/\sqrt{n}} = z_{1-\frac{\epsilon}{2}}$$

که $z_{1-\frac{\epsilon}{2}}$ طوی لن توزیع $N(0, 1)$ است

که سطح زیر منحنی توزیع از $-\infty$ تا $z_{1-\frac{\epsilon}{2}}$ مساوی $1 - \frac{\epsilon}{2}$ است:

بازده اعتماد - ۳

(به عنوان مثال برای $\epsilon = 0.05$ این طول برابر با ۱.۹۶ و به ازای $\epsilon = 0.01$ این طول برای ۲.۵۷۵ می باشد)

بنابراین: $\mu_1 = \mu - z_{1-\frac{\epsilon}{2}} \frac{\sigma}{\sqrt{n}}$ و $\mu_2 = \mu + z_{1-\frac{\epsilon}{2}} \frac{\sigma}{\sqrt{n}}$

و در نتیجه: $P(\mu - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\epsilon}{2}} < \hat{\mu} < \mu + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\epsilon}{2}}) = 1 - \epsilon$
 حاصل نابرابری های جدول و انتگرال روی μ :

$$P(\hat{\mu} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\epsilon}{2}} < \mu < \hat{\mu} + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\epsilon}{2}}) = 1 - \epsilon$$

پس: $L = \hat{\mu} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\epsilon}{2}}$, $\bar{L} = \hat{\mu} + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\epsilon}{2}}$

$$= \bar{x} - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\epsilon}{2}} \quad = \bar{x} + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\epsilon}{2}}$$

مثال ۱- از جامه های با توزیع نرمال با واریانس $\sigma^2 = 100$ ، آمار زیر استخراج شده است:

{۶۵، ۴۵، ۴۳، ۴۰، ۶۴، ۵۸، ۵۲، ۵۶، ۶۲}

بازده اعتماد ۹۵٪ می بایست جامه را تعیین کنند: حل: $n=9$ ، $\epsilon=0.05$

$$\hat{\mu} = \bar{x} = \frac{1}{9}(62+56+52+58+64+40+43+45+65) = 54$$

$$z_{1-\frac{\epsilon}{2}} = z_{0.975} = 1.96$$

$$L = \bar{x} - \frac{10}{\sqrt{9}} \times 1.96$$

$$= 54 - \frac{10}{3} \times 1.96 = 54 - 6.53 = 47.47$$

$$\bar{L} = \bar{x} - \frac{10}{\sqrt{9}} \times 1.96 = 54 + 6.53 = 60.53$$

یعنی با اقبال ۹۵٪: $\mu \in [47.47, 60.53]$

نم- برای توزیع پیوسته با تابع چگالی اقبال: $f_x(x)$ اگر داشته

با سهم مثبت: $\int_a^b f_x(x) dx = P = A < 1$ ، $b-a$ وقتی می نمیم است که $f_x(a) = f_x(b)$ باشد.

بازه اعتماد - ۴
اثبات : با یکاربری ضریب نامعین لاگرانژ :

$$I = b - a + \lambda \left(\int_a^b f_x(x) dx - A \right)$$

$$\frac{\partial I}{\partial b} = 1 + \lambda \frac{\partial}{\partial b} \int_a^b f_x(x) dx = 1 + \lambda f_x(b) = 0$$

$$\frac{\partial I}{\partial a} = -1 + \lambda \frac{\partial}{\partial a} \int_a^b f_x(x) dx = -1 - \lambda f_x(a) = 0$$

از جمع این دو معادله بدست می آوریم :
زیرا به ازای $\lambda = 0$ ، دو معادله بالا بنابر $\lambda(f_x(b) - f_x(a)) = 0$ ، $\lambda \neq 0$ است
لذا $f_x(b) - f_x(a) = 0$ و $f_x(b) = f_x(a)$ یعنی $f_x(b) = f_x(a)$

توجه : همانطور که قبلاً ذکر شد در مورد توزیع های متقارن، شرط $f_x(b) = f_x(a)$

معادل است با اینکه سطح زیر نمودار تابع چگالی احتمال در طرفین بازه $[a, b]$ با هم برابر باشد و اگر $A = 1 - \epsilon$ باشد این ساحت که هر کدام مساوی $\frac{\epsilon}{2}$ است این نکته را قوی و قوی که توزیع متقارن نباشد بصورت تقریب اعمال می کنیم

۲- روش کلی برای تعیین بازه اعتماد (۱-۴) ۱۰۰ درصد پارامتر توزیع .

فرض کنیم که برای n گوی از روشهای برآورد ، پارامتر α به صورت :

$(x_1, x_2, \dots, x_n)$ برآورد شده است . ابتداء اعداد α_1 و α_2 ($\alpha_1 < \alpha_2$)

را طوری تعیین می کنیم که $P(\alpha_1 < \hat{\alpha} < \alpha_2) = 1 - \epsilon$ و $\alpha_1 - \alpha_2$ می نامیم

باشد با توجه به لم بالا و نکته ذکر شده در صورت آن ، بدست می آوریم

$$P(\hat{\alpha} < \alpha_1) = \frac{\epsilon}{2} \text{ و } P(\hat{\alpha} > \alpha_2) = \frac{\epsilon}{2} \text{ با استفاده از توزیع}$$

$\hat{\alpha}$ ، α_1 و α_2 به صورت کولمبی از α و ϵ بدست می آید .

فرض کنیم $\alpha_1 = \alpha_1(\alpha, \epsilon)$ و $\alpha_2 = \alpha_2(\alpha, \epsilon)$. لذا داریم :

$$P(\alpha_1(\alpha, \epsilon) < \hat{\alpha} < \alpha_2(\alpha, \epsilon)) = 1 - \epsilon$$

می آوریم $P(\alpha_1(\hat{\alpha}, \epsilon) < \alpha < \alpha_2(\hat{\alpha}, \epsilon)) = 1 - \epsilon$ به این ترتیب حدود بازه اعتماد بدست می آید

بازة اعتماد - ۵

سوال ۲- میخواهم بازة اعتماد (۴-۱-۱۰۰ درصد) برای تخمین میانگین μ داشته باشم. فرض کنیم $N(\mu, \sigma^2)$ باشد. برای آنکه برآورد $\hat{\mu} = \bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ و $\hat{\sigma}^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$ درست آمده است

ابتدا μ_1 و μ_2 را به طوری که $P(\mu_1 \leq \bar{X} \leq \mu_2) = 1 - \epsilon$ باشد تعیین می‌کنیم و برای اینجا ریشه توزیع $\bar{X}$ می‌رویم. ولی در این معیار $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$ دارای توزیع $N(0,1)$ نیست بلکه دارای توزیع t -استودنت با $\nu = n-1$ درجه آزادی است:

لذا:

$$P\left(\frac{\mu_1 - \mu}{\sigma/\sqrt{n}} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{\mu_2 - \mu}{\sigma/\sqrt{n}}\right) = 1 - \epsilon$$

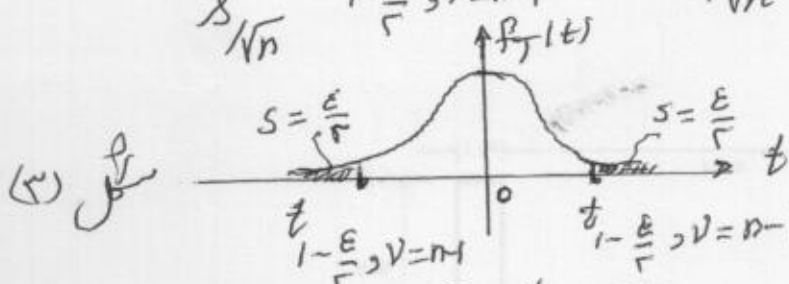
لذا باید

$$\frac{\mu_2 - \mu}{\sigma/\sqrt{n}} = t_{1-\frac{\epsilon}{2}, \nu=n-1} \quad \text{و} \quad \frac{\mu_1 - \mu}{\sigma/\sqrt{n}} = -t_{1-\frac{\epsilon}{2}, \nu=n-1}$$

و از اینجا:

$$\mu_1 = \mu - \frac{\sigma}{\sqrt{n}} t_{1-\frac{\epsilon}{2}, \nu=n-1}$$

$$\mu_2 = \mu + \frac{\sigma}{\sqrt{n}} t_{1-\frac{\epsilon}{2}, \nu=n-1}$$



مقدار تابع چگالی لگال توزیع t -استودنت

توجه: $t_{1-\frac{\epsilon}{2}, \nu=n-1}$ طولی از نمودار

تابع چگالی لگال t -استودنت با درجه آزادی $\nu = n-1$ است که سطح زیر آنی از $-\infty$ تا t برابر $1 - \frac{\epsilon}{2}$ است: پس داریم:

$$P(\mu_1 \leq \bar{X} \leq \mu_2) = 1 - \epsilon \Rightarrow P\left(\mu - \frac{\sigma}{\sqrt{n}} t_{1-\frac{\epsilon}{2}, \nu=n-1} \leq \bar{X} \leq \mu + \frac{\sigma}{\sqrt{n}} t_{1-\frac{\epsilon}{2}, \nu=n-1}\right) = 1 - \epsilon$$

پس حاصل تا برآورد های داخل برآورد روی μ درست می‌آید

$$L_1 = \bar{X} - \frac{\sigma}{\sqrt{n}} t_{1-\frac{\epsilon}{2}, \nu=n-1}$$

$$L_2 = \bar{X} + \frac{\sigma}{\sqrt{n}} t_{1-\frac{\epsilon}{2}, \nu=n-1}$$

$$P(L_1 \leq \mu \leq L_2) = 1 - \epsilon$$

مسئله عددی - همان مفروضات مثال (۱) با این تفاوت که ^{جامعه} μ را نمی دانیم و برآورد آمار باید برآورد شود

حل: $\{45, 43, 40, 44, 58, 52, 56, 42\} \rightarrow$ آمار

$$\Rightarrow \bar{x} = \frac{\sum x_i}{n} = 54, \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n (x_i - 54)^2 = 724, s^2 = \frac{724}{9-1} = 90.5$$

واریانس جدول t - استودنت برای درجه آزادی $\nu = 8$ و $\alpha = 0.05$

$t_{0.975, \nu=8} = 2.306$ استخراج می شود.

$$\Rightarrow \Delta = \sqrt{90.5} \Rightarrow \frac{\Delta}{\sqrt{n}} = \frac{9.51215}{\sqrt{9}} = 3.171$$

$$L = \bar{x} - \frac{\Delta}{\sqrt{n}} t = 54 - 3.171 \times 2.306 = 46.69$$

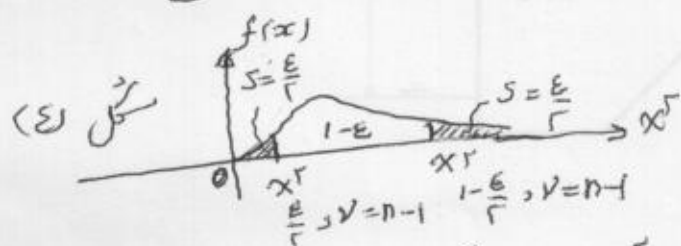
$$U = \bar{x} + \frac{\Delta}{\sqrt{n}} t = 54 + 3.171 \times 2.306 = 61.31$$

مثال ۳ - بازه اعداد واریانس توزیع نرمال - فرض کنید μ و σ از توزیع $N(\mu, \sigma)$ مجهول هستند و آنها را از روی آمار استخراجی از جامعه برآورد کرده ایم:

بطوریکه $\hat{\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ و $\hat{\sigma}^2 = s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$ می دانیم که

$\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{\sigma^2}$ دارای توزیع χ^2 با درجه آزادی $\nu = n-1$ است

لذا طبق شکل (ع):



$$P\left(\chi_{\frac{\epsilon}{2}, \nu}^2 \leq \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sigma^2} \leq \chi_{1-\frac{\epsilon}{2}, \nu}^2\right) = 1-\epsilon$$

طرفین نامرئی در یک طرف قرار

$$\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sigma^2} = \frac{s^2}{\sigma^2} \times \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{s^2}$$

$$\Rightarrow P\left(\frac{\chi_{1-\frac{\epsilon}{2}, \nu}^2}{\chi_{\frac{\epsilon}{2}, \nu}^2} \leq \frac{s^2}{\sigma^2} \leq \frac{\chi_{1-\frac{\epsilon}{2}, \nu}^2}{\chi_{\frac{\epsilon}{2}, \nu}^2}\right) = 1-\epsilon$$

مثال عددی $n=99$ و $\epsilon=0.01$ و $\bar{x}=97500$ و $s^2=21000$

$\chi_{0.005, 98}^2 = 124.18$ و $\chi_{0.995, 98}^2 = 47.4$

$L = 97500 \times 1.06$ و $U = 97500 \times 1.06$

مثال (۳) را وقتی که μ را می دانیم و σ را نمی دانیم حل کنید.

آزمون کای‌سکوئر برازش (Goodness of Fit Test) - ۱

طرح مسئله: فرض کنید توزیع آماری جامعه‌ای (مستقیم یا نامستقیم) معلوم نیست و برای مشاهده $\{X_1, X_2, X_3, \dots, X_n\}$ از این جامعه نمونه‌گیری کرده‌ایم و دارای توزیع معینی است (مثلاً با رسم هیستوگرام این آمار و تقایب آن با نمودار تابع فراوانی توزیع پیش‌بینی‌شده). این آزمون به دو طریق سیر است: ۱- آزمون χ^2 ، ۲- آزمون کولموگوروف - اسمیرنوف.

I - آزمون χ^2 (Chi-square Goodness of Fit Test)

برای روش ملاک آزمون عدد $\chi^2 = \sum_{i=1}^k \frac{(F_i - F'_i)^2}{F'_i}$ است، که در این فرمول F_i که فراوانی‌های مشاهده شده و F'_i فراوانی‌های نظری (محاسبه شده بر اساس توزیع پیش‌بینی‌شده می‌شود) برای هر طبقه سبزی آمار به دست آمده می‌باشد.

اگر دانسته باشیم X ، بازه $[a, b]$ باشد، این بازه را به k زیر بازه‌ها تقسیم می‌کنیم. در این صورت F_i برابر است با فراوانی (تعداد) افراد مشاهده شده در بازه i ، و F'_i نیز فراوانی نظری (تعداد محاسبه شده) است که از روی توزیع که قرار است به داده‌ها برازش داده شود محاسبه می‌شود و داریم:

$$\sum_{i=1}^k F_i = \sum_{i=1}^k F'_i = n \quad (2)$$

توجه: اگر F_i طبقه‌ای کوچکتر از F'_i باشد، آن طبقه را با طبقه مجاور سرهم می‌کنیم. نکته دیگر اینکه ممکن است درباره مقدار پارامتر (یا پارامترهای) مربوط به توزیع پیش‌بینی‌شده، فرضی در نظر گرفته شود و یا اینکه ضریب و مقدار آنها را از روی آمار استخراجی برآورد کرده باشیم.

نتیجه آزمون از مقایسه χ^2 با χ^2_{α} مربوط به سطح احتمال ۱- α (مثلاً ۰.۰۵ یا ۰.۰۱) با درجه آزادی $v = k - p - 1$ که k تعداد طبقات، p تعداد پارامترهای برآورد شده توزیع پیش‌بینی‌شده، صورت می‌گیرد، به طوریکه اگر $\chi^2_{\alpha} > \chi^2$ بود، آنگاه فرض اینکه جامعه فریبور دارای توزیع پیش‌بینی‌شده است، رد می‌شود و در غیر این صورت رد نمی‌شود.

اگر $\chi^2 > \chi^2_{\alpha}$ بود ملاک (۱) چندان معییر نیست و جای آن از عبارت: $\chi^2 = \sum_{i=1}^k \frac{(F_i - F'_i)^2}{F_i}$ استفاده می‌کنیم.

آزمون کولموگوروف برای ۲-
 مثال ۱- در منطقه معینی در طی ۱۰۰ سال تعداد سیلابهای سالانه زیر ثبت
 شده است.

تعداد سیلابها	۰	۱	۲	۳	۴	۵	۶
تعداد سال وقوع	۵۲	۲۸	۱۲	۵	۲	۱	۰

الفنی در ۵۲ سال سیلابی نداشته ایم، در ۲۸ سال در هر سال یک سیلاب داشته ایم و ۱۰۰۰

سوال آیا تعداد سیلابها در هر سال از توزیع پواسن تبعیت می کند؟

حل - فرض H_0 تعداد سیلابها در سال دارای توزیع پواسن است. با فرض قبول H_0 با استفاده از داده های فوق پارامتر مربوط به این توزیع را برآورد می کنیم

$$\lambda = \bar{X} = \frac{52 \times 0 + 28 \times 1 + 12 \times 2 + 5 \times 3 + 2 \times 4 + 1 \times 5}{100} = \frac{10}{100} = 0.1$$

لذا:

$$P_x(x) = e^{-\lambda} \frac{\lambda^x}{x!} = e^{-0.1} \times \frac{(0.1)^x}{x!}$$

$x=0, 1, 2, 3, \dots$

حال جدول مربوط به فراوانیهای نظری و همچنین فراوانی های مشاهده شده را تکمیل می دهیم:

x_i	$P(x_i)$	$F'_i = 100 \cdot P(x_i)$	F_i	F'_i	F_i
۰	$e^{-0.1} = 0.4493$	۴۴,۹۳	۵۲	۴۴,۹۳	۵۲
۱	$e^{-0.1} \times \frac{0.1}{1!} = 0.2595$	۲۵,۹۵	۲۸	۲۵,۹۵	۲۸
۲	$e^{-0.1} \times \frac{(0.1)^2}{2!} = 0.1238$	۱۲,۳۸	۱۲	۱۲,۳۸	۱۲
۳	$e^{-0.1} \times \frac{(0.1)^3}{3!} = 0.0373$	۳,۸۳	۵	۴,۷۲	۸
۴	$e^{-0.1} \times \frac{(0.1)^4}{4!} = 0.0077$	۰,۷۷	۲	سرهم کردن دسته های با $F' < 5$	
۵	$e^{-0.1} \times \frac{(0.1)^5}{5!} = 0.0012$	۰,۱۲	۱		

$$\chi^2 = \sum_{i=1}^4 \frac{(F'_i - F_i)^2}{F'_i} = \frac{(44.93 - 52)^2}{44.93} + \frac{(25.95 - 28)^2}{25.95} + \frac{(12.38 - 12)^2}{12.38} + \frac{(4.72 - 8)^2}{4.72}$$

آزمون نیکوتس برابرش - ۳

$\chi^2 = 1,1125 + 1,7580 + 0,3939 + 2,2792 = 5,5437$
 با توجه به اینکه تعداد طبقه ها $k=4$ و یک پارامتر از توزیع برآورده شده است، لذا
 $v = k - 1 - 1 = 2$ و از روی جدول توزیع χ^2 در سطح معنی دار شدن ۵٪ و برای درجه آزادی
 $v=2$ عدد $\chi^2_{0.05, v=2} = 5,99$ استخراج می شود. با توجه به اینکه $5,5437 < 5,99$
 است. لذا فرض H_0 رد نمی شود.

مثال ۲- از حیواناتی ۱۰۰ فرد نمونه گیری کرده ایم براساس طبقه بندی زیر

طبقه	F_i
۶۰-۶۲	۵
۶۳-۶۵	۱۸
۶۶-۶۸	۴۲
۶۹-۷۱	۲۷
۷۲-۷۴	۸

فراوانیهای مشاهده شده چنین اند
 آزمون نرمال بودن توزیع جامعه را انجام دهید
 حل - جامعه دارای توزیع نرمال است: H_0
 با قبول فرض H_0 ، میانگین و واریانس جامعه را
 برآورد می کنیم

$$\hat{\mu} = \bar{X} = \frac{5 \times 61 + 18 \times 64 + 42 \times 67 + 27 \times 70 + 8 \times 73}{100} = \frac{9745}{100} = 97,45$$

$$\hat{\sigma}^2 = s^2 = \frac{5(61-97,45)^2 + 18(64-97,45)^2 + 42(67-97,45)^2 + 27(70-97,45)^2 + 8(73-97,45)^2}{99}$$

$$= 8,526 \Rightarrow \hat{\sigma} = 2,92$$

طبقه	کراتهای طبقه	کراتهای تصحیح z	سطح زیر انتی توزیع z از $N(0,1)$	مساحت مربوط به هر طبقه	فراوانیهای نظری F_i	فراوانیهای مشاهده شده F_i
۶۰-۶۲	۵۹,۵	-۲,۷۲	۰,۴۹۹۷	۰,۰۴۱۳	۴,۱۳	۵
۶۳-۶۵	۶۲,۵	-۱,۷۰	۰,۴۵۵۴	۰,۲۰۹۸	۲۰,۹۸	۱۸
۶۶-۶۸	۶۵,۵	-۰,۹۷	۰,۲۴۸۹	۰,۳۸۹۲	۳۸,۹۲	۴۲
۶۹-۷۱	۶۸,۵	۰,۳۶	۰,۱۴۳۶	۰,۲۷۷۱	۲۷,۷۱	۲۷
۷۲-۷۴	۷۱,۵	۱,۳۹	۰,۴۱۷۷	۰,۰۷۴۲	۷,۴۲	۸
	۷۴,۵	۲,۴۱	۰,۴۹۲۰			

آزمون نیکوتی برابرش - ۴

(توضیح: اعداد مربوط به ستون سوم: کمرانهای پیفر ۲، به این ترتیب مرتب آمده اند)

$$\Rightarrow Z_1 = \frac{X_1 - \bar{X}}{s} = \frac{59.5 - 67.45}{2.92} = -2.72, \quad Z_2 = \frac{X_2 - \bar{X}}{s} = \frac{92.5 - 67.45}{2.92} = 8.58$$

حالت: (و...)

$$\chi^2 = \frac{(5 - 4.13)^2}{4.13} + \frac{(11 - 2.16)^2}{2.16} + \frac{(42 - 38.92)^2}{38.92} + \frac{(27 - 27.71)^2}{27.71} + \frac{(1 - 7.43)^2}{7.43}$$

برای این آزمون $h = 5$ ، تعداد پارامترهای برآورد شده $p = 2$ ، لذا $df = 3$
 و از روی جدول توزیع χ^2 ، بدست می آوریم $\chi^2_{0.95, 3} = 7.88$ ، در این مثال $\chi^2 = 5.99 < 7.88$ و فرض H_0 رد نمی شود

مثال ۳- از جامعه ای که افراد آن دارای یکی از دو خصوصیت A و $B = \bar{A}$ می باشد ۱۰۰ نفر را به تصادف انتخاب کرده ایم ۳۹ نفر دارای خصوصیت A و ۶۱ نفر دارای خصوصیت B هستند. آیا می توان پذیرفت که فراوانی نسبی دو گروه در جامعه با هم برابرند حل:

$H_0: \frac{1}{2}$ افراد جامعه دارای خصوصیت A و $\frac{1}{2}$ دیگر دارای خصوصیت B هستند
 جدول مربوط به فراوانیهای مشاهده شده و فراوانی های نظری چنین است:

F_i	F'_i	$F_i - F'_i$	$ F_i - F'_i - 0.5$
۳۹	۵۰	-۱۱	۱۰.۵
۶۱	۵۰	۱۱	۱۰.۵

$$\chi^2 = \sum \frac{(F_i - F'_i - 0.5)^2}{F'_i} = \frac{(10.5)^2}{50} + \frac{(10.5)^2}{50} = 4.41$$

و از روی جدول توزیع χ^2 ، $\chi^2_{0.95, 1} = 3.84$ و از آنکه $\chi^2 = 4.41 > 3.84$ ، نتیجه می شود که فرض H_0 رد می شود و نمی توان پذیرفت که فراوانی دو گروه در جامعه یکسان باشد و ۹۵٪ به تفاوت خود اعتماد داریم و ۵٪ احتمال اشتباه در این قضاوت می رود

سؤال مختلف آمار ۱ -

۱- برای توزیع $N(2, 9)$ ، مطلوب است محاسبه α به طوریکه:

$$P(X \leq \alpha) = 2P(X \geq \alpha) \quad \text{جواب: } \alpha = 3.29$$

۲- متغیرهای $X_1, X_2, \dots, X_n$ که مستقل از یکدیگر توزیع شده اند به طوریکه $X_j \sim N(\mu_j, \sigma^2)$ ($j=1, 2, \dots, n$) با فرض:

$$Y = [(X_1 - \mu_1)^2 + (X_2 - \mu_2)^2 + \dots + (X_n - \mu_n)^2]^{\frac{1}{2}}$$

داریم: $E(Y) = \sqrt{2} \sigma \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})}$

۳- X و Y دارای توزیع بی نامیه با توابع احتمال:

$$p_X(x) = P(X=x) = p(1-p)^x \quad x=0, 1, 2, 3, \dots$$

$$p_Y(y) = P(Y=y) = q(1-q)^y \quad y=0, 1, 2, 3, \dots$$

می باشند که $0 < p, q < 1$

جواب: توزیع متغیر $Z = X + Y$ را تعیین کنید

$$p_Z(z) = P(Z=z) = \frac{pq}{q-p} \left((1-p)^{z+1} - (1-q)^{z+1} \right) \quad p \neq q$$

و وقتی که $p=q$ ، آنگاه

$$p_Z(z) = (z+1)p^2(1-p)^z$$

۴- نشان دهید میانگین و واریانس توزیع بتا با تابع چگالی احتمال:

$$f_X(x) = \frac{1}{B(p, q)} x^{p-1} (1-x)^{q-1} \quad (0 < x < 1)$$

$$B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \quad \text{که } \Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$$

$$\mu = \frac{p}{p+q} \quad \text{و} \quad \sigma^2 = \frac{pq}{(p+q)^2(p+q+1)}$$

وقتی $p=q$ است، ضریب چوگلی توزیع صفر است. ب- با یکبارگیری

روش گسترده را برآورد پارامترهای p و q را بر حسب $\bar{x} = \frac{1}{n} \sum_{i=1}^n X_i$

$$\delta^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{x})^2 \quad \text{و} \quad \text{جواب:}$$

$$\hat{p} = \bar{x} \left(\frac{\bar{x}(1-\bar{x})}{\delta^2} - 1 \right), \quad \hat{q} = (1-\bar{x}) \left(\frac{\bar{x}(1-\bar{x})}{\delta^2} - 1 \right)$$

سؤال مختلف کار ۲-

۵- برای توزیع کنواخت $u(a, b)$ تابع چگالی احتمال:

$$f_X(x) = \frac{1}{b-a} \quad a < x < b$$

در غیر این صورت = 0

$$\mu_X = \frac{a+b}{2} \quad \text{و} \quad \sigma_X^2 = \frac{(b-a)^2}{12}$$

ب- اگر $\{X_1, X_2, \dots, X_n\}$ یک نمونه گزینی به حجم n از این توزیع باشد:

$$E(\min\{X_1, X_2, \dots, X_n\}) \quad \text{و} \quad E(\max\{X_1, X_2, \dots, X_n\})$$

را حساب کنید. جواب

$$E(\max\{X_1, \dots, X_n\}) = \frac{nb+a}{n+1}$$

$$E(\min\{X_1, \dots, X_n\}) = \frac{na+b}{n+1}$$

۶- اگر X دارای توزیع پواسنه با تابع چگالی $f_X(x)$ باشد ثابت کنید تغییر $z = F_X(x)$ دارای توزیع کنواخت $u(0, 1)$ است.

نشان دهید چگونه با داشتن جدول اعداد تصادفی (که دارای توزیع $u(0, 1)$ است)، می توان داده گزینی با توزیع های گمبل و وایبول ساخت.

اثبات تست کول: با فرض $z = F_X(x)$ تغییرات z از ۰ تا ۱ است

فرض کنید $f_2(z)$ تابع چگالی احتمال متغیر z باشد. داریم:

$$\begin{aligned} f_2(z) dz &= P(z \leq Z \leq z + dz) = P(z \leq F_X(X) \leq z + dz) \\ &= P(F_X^{-1}(z) \leq X \leq F_X^{-1}(z + dz)) = \int_{F_X^{-1}(z)}^{F_X^{-1}(z + dz)} f_X(x) dx \\ &= F_X(F_X^{-1}(z + dz)) - F_X(F_X^{-1}(z)) = z + dz - z = dz \end{aligned}$$

$$f_2(z) dz = dz \Rightarrow \underline{f_2(z) = 1} \quad 0 < z < 1$$

۷- نشان دهید اگر $\hat{\alpha} = \phi(X_1, X_2, \dots, X_n)$ یک برآورد ناریب پارامتر α باشد، $\hat{\alpha}^2$ برآورد ناریب پارامتر α^2 است یعنی داریم

$$E(\hat{\alpha}^2) \neq \alpha^2$$

مسائل مختلف آمار ۳-

۱- برای توزیع پواسن به تابع چگالی (فصل): $\lambda > 0, f_X(x) = \frac{\lambda}{x!} e^{-\lambda|x|}$

پارامتر λ از روی نمونه $\{x_1, x_2, \dots, x_n\}$ برآورد کنیم.

جواب: $\hat{\lambda} = \frac{n}{\sum_{i=1}^n |x_i|}$

۲- خواسته شده قبل از برای توزیع پواسن: $\lambda > 0, f_X(x) = \lambda x^{\lambda-1}$ (درست آوردید)

جواب: روش حداکثر تابع درست نمایی

$\hat{\lambda} = -\frac{n}{\sum_{i=1}^n \ln x_i}$ روش گسترده ها:

۱۰- نشان دهید اگر X دارای توزیع گاما با پارامترهای λ و η باشد
 $f_X(x) = \frac{\lambda^\eta}{\Gamma(\eta)} x^{\eta-1} e^{-\lambda x}$ متغیر $Y = 2\lambda X$ دارای توزیع آلفا باشد
 اگر $\eta = 2$ است.

۱۱- نشان دهید اگر متغیرهای X و Y دارای توزیع پواسن به ترتیب با پارامترهای λ و μ باشند، متغیر $Z = X + Y$ دارای توزیع پواسن با پارامتر $\lambda + \mu$ است.

۱۲- اگر متغیرهای مستقل X و Y به ترتیب دارای توزیع های گاما:

$f_X(x) = \frac{\lambda_1^{\eta_1}}{\Gamma(\eta_1)} x^{\eta_1-1} e^{-\lambda_1 x}$ ، $f_Y(y) = \frac{\lambda_2^{\eta_2}}{\Gamma(\eta_2)} y^{\eta_2-1} e^{-\lambda_2 y}$ باشند متغیر

$Z = X + Y$ دارای توزیع گاما با تابع چگالی (فصل):

$f_Z(z) = \frac{\lambda_1^{\eta_1} \lambda_2^{\eta_2}}{\Gamma(\eta_1 + \eta_2)} z^{\eta_1 + \eta_2 - 1} e^{-(\lambda_1 + \lambda_2)z}$ است.

سؤال مختلف آمار - ۴

۱۳- فرض کنید در طول ماه اول میانگین و انحراف معیار بارندگی ماهانه به ترتیب

۰/۷۵ و ۰/۴۳۳ باشد. همچنین میانگین و انحراف معیار

بارندگی در طول ماه دوم به ترتیب ۳۰۰۰ و ۱۸۹۹ باشد

با فرض اینکه تعداد بارندگی ماهانه توسط توزیع گاما تقریب زده شود

و شراک بارندگی ماه دوم مستقل از ماه اول باشد، احتمال اینکه در طول

این دو ماه بیش از ۳ اینچ بارندگی داشته باشیم چقدر است؟

(جواب: $p = 1 - 0.772 = 0.228$) ابتدا با داشتن میانگین و واریانس دو توزیع

پارامترهای دو توزیع یعنی λ و η را محاسبه می کنیم دیدیم که برای توزیع گاما با پارامترهای

اولی، داریم $\mu = \frac{\eta}{\lambda}$ و $\sigma^2 = \frac{\eta}{\lambda^2}$ پس برای توزیع بارندگی ماه اول:

$$\Rightarrow \frac{\eta_1}{\lambda_1} = 0.75 \text{ و } \frac{\eta_1}{\lambda_1^2} = (0.433)^2 \Rightarrow \eta_1 = 3 \text{ و } \lambda_1 = 4$$

و برای توزیع بارندگی ماه دوم:

$$\Rightarrow \frac{\eta_2}{\lambda_2} = 3000 \text{ و } \frac{\eta_2}{\lambda_2^2} = (1899)^2 \Rightarrow \eta_2 = 12 \text{ و } \lambda_2 = 4$$

پس مجموع بارندگی طی دو ماه دارای توزیع گاما با پارامترهای

$$\lambda = \lambda_1 = \lambda_2 = 4 \text{ و } \eta = \eta_1 + \eta_2 = 15 \text{ است و برای این توزیع محاسبه}$$

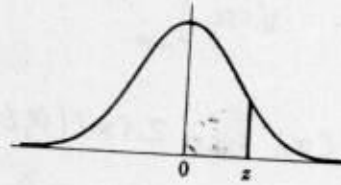
$$P(Z > 3) = 1 - P(Z < 3) \text{ ساده است.}$$

۱۴- یک کاس را ۱۲۰ بار پرتاب کرده ایم تعداد دفعات ظهور هر یک از اعداد ۱ تا ۶ طبق جدول زیر است. در سطح معنی دار بودن ۵٪، فرض سالم بودن کاس را آزمون کنید.

عدد	۱	۲	۳	۴	۵	۶
F_i	۲۵	۱۷	۱۵	۲۳	۲۴	۱۶

$\Rightarrow \chi^2 = 5.00$

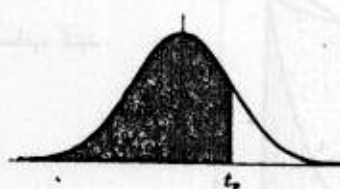
**AREAS
under the
STANDARD
NORMAL CURVE
from 0 to z**



z	0	1	2	3	4	5	6	7	8	9
0-0	0-0000	0-0040	0-0080	0-0120	0-0160	0-0199	0-0239	0-0279	0-0319	0-0359
0-1	0-0398	0-0438	0-0478	0-0517	0-0557	0-0596	0-0636	0-0675	0-0714	0-0754
0-2	0-0793	0-0832	0-0871	0-0910	0-0948	0-0987	0-1026	0-1064	0-1103	0-1141
0-3	0-1179	0-1217	0-1255	0-1293	0-1331	0-1368	0-1406	0-1443	0-1480	0-1517
0-4	0-1554	0-1591	0-1628	0-1664	0-1700	0-1736	0-1772	0-1808	0-1844	0-1879
0-5	0-1915	0-1950	0-1985	0-2019	0-2054	0-2088	0-2123	0-2157	0-2190	0-2224
0-6	0-2258	0-2291	0-2324	0-2357	0-2389	0-2422	0-2454	0-2486	0-2518	0-2549
0-7	0-2580	0-2612	0-2642	0-2673	0-2704	0-2734	0-2764	0-2794	0-2823	0-2852
0-8	0-2881	0-2910	0-2939	0-2967	0-2996	0-3023	0-3051	0-3078	0-3106	0-3133
0-9	0-3159	0-3186	0-3212	0-3238	0-3264	0-3289	0-3315	0-3340	0-3365	0-3389
1-0	0-3413	0-3438	0-3461	0-3485	0-3508	0-3531	0-3554	0-3577	0-3599	0-3621
1-1	0-3643	0-3665	0-3686	0-3708	0-3729	0-3749	0-3770	0-3790	0-3810	0-3830
1-2	0-3849	0-3869	0-3888	0-3907	0-3925	0-3944	0-3962	0-3980	0-3997	0-4015
1-3	0-4032	0-4049	0-4066	0-4082	0-4099	0-4115	0-4131	0-4147	0-4162	0-4177*
1-4	0-4192	0-4207	0-4222	0-4236	0-4251	0-4265	0-4279	0-4292	0-4306	0-4319
1-5	0-4332	0-4345	0-4357	0-4370	0-4382	0-4394	0-4406	0-4418	0-4429	0-4441
1-6	0-4452	0-4463	0-4474	0-4484	0-4495	0-4505	0-4515	0-4525	0-4535	0-4545
1-7	0-4554	0-4564	0-4573	0-4582	0-4591	0-4599	0-4608	0-4616	0-4625	0-4633
1-8	0-4641	0-4649	0-4656	0-4664	0-4671	0-4678	0-4686	0-4693	0-4699	0-4706
1-9	0-4713	0-4719	0-4726	0-4732	0-4738	0-4744	0-4750	0-4756	0-4761	0-4767
2-0	0-4772	0-4778	0-4783	0-4788	0-4793	0-4798	0-4803	0-4808	0-4812	0-4817
2-1	0-4821	0-4826	0-4830	0-4834	0-4838	0-4842	0-4846	0-4850	0-4854	0-4857
2-2	0-4861	0-4864	0-4868	0-4871	0-4875	0-4878	0-4881	0-4884	0-4887	0-4890
2-3	0-4893	0-4896	0-4898	0-4901	0-4904	0-4906	0-4909	0-4911	0-4913	0-4916
2-4	0-4918	0-4920	0-4922	0-4925	0-4927	0-4929	0-4931	0-4932	0-4934	0-4936
2-5	0-4938	0-4940	0-4941	0-4943	0-4945	0-4946	0-4948	0-4949	0-4951	0-4952
2-6	0-4953	0-4955	0-4956	0-4957	0-4959	0-4960	0-4961	0-4962	0-4963	0-4964
2-7	0-4965	0-4966	0-4967	0-4968	0-4969	0-4970	0-4971	0-4972	0-4973	0-4974
2-8	0-4974	0-4975	0-4976	0-4977	0-4977	0-4978	0-4979	0-4979	0-4980	0-4981
2-9	0-4981	0-4982	0-4982	0-4983	0-4984	0-4984	0-4985	0-4985	0-4986	0-4986
3-0	0-4987	0-4987	0-4987	0-4988	0-4988	0-4989	0-4989	0-4989	0-4990	0-4990
3-1	0-4990	0-4991	0-4991	0-4991	0-4992	0-4992	0-4992	0-4992	0-4993	0-4993
3-2	0-4993	0-4993	0-4994	0-4994	0-4994	0-4994	0-4994	0-4995	0-4995	0-4995
3-3	0-4995	0-4995	0-4995	0-4996	0-4996	0-4996	0-4996	0-4996	0-4996	0-4997
3-4	0-4997	0-4997	0-4997	0-4997	0-4997	0-4997	0-4997	0-4997	0-4997	0-4998
3-5	0-4998	0-4998	0-4998	0-4998	0-4998	0-4998	0-4998	0-4998	0-4998	0-4998
3-6	0-4998	0-4998	0-4999	0-4999	0-4999	0-4999	0-4999	0-4999	0-4999	0-4999
3-7	0-4999	0-4999	0-4999	0-4999	0-4999	0-4999	0-4999	0-4999	0-4999	0-4999
3-8	0-4999	0-4999	0-4999	0-4999	0-4999	0-4999	0-4999	0-4999	0-4999	0-4999
3-9	0-5000	0-5000	0-5000	0-5000	0-5000	0-5000	0-5000	0-5000	0-5000	0-5000

Appendix III

PERCENTILE VALUES (t_p)
for
STUDENT'S t DISTRIBUTION
with v degrees of freedom
(shaded area = p)

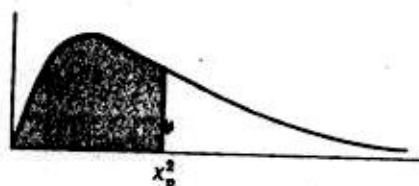


v	$t_{0.995}$	$t_{0.99}$	$t_{0.975}$	$t_{0.95}$	$t_{0.90}$	$t_{0.80}$	$t_{0.75}$	$t_{0.70}$	$t_{0.60}$	$t_{0.55}$
1	63.66	31.82	12.71	6.31	3.08	1.376	1.000	0.727	0.325	0.158
2	9.92	6.96	4.30	2.92	1.89	1.061	0.816	0.617	0.289	0.142
3	5.84	4.54	3.18	2.35	1.64	0.978	0.765	0.584	0.277	0.137
4	4.60	3.75	2.78	2.13	1.53	0.941	0.741	0.569	0.271	0.134
5	4.03	3.36	2.57	2.02	1.48	0.920	0.727	0.559	0.267	0.132
6	3.71	3.14	2.45	1.94	1.44	0.906	0.718	0.553	0.265	0.131
7	3.50	3.00	2.36	1.90	1.42	0.896	0.711	0.549	0.263	0.130
8	3.36	2.90	2.31	1.86	1.40	0.889	0.706	0.546	0.262	0.130
9	3.25	2.82	2.26	1.83	1.38	0.883	0.703	0.543	0.261	0.129
10	3.17	2.76	2.23	1.81	1.37	0.879	0.700	0.542	0.260	0.129
11	3.11	2.72	2.20	1.80	1.36	0.876	0.697	0.540	0.260	0.129
12	3.06	2.68	2.18	1.78	1.36	0.873	0.695	0.539	0.259	0.128
13	3.01	2.65	2.16	1.77	1.35	0.870	0.694	0.538	0.259	0.128
14	2.98	2.62	2.14	1.76	1.34	0.868	0.692	0.537	0.258	0.128
15	2.95	2.60	2.13	1.75	1.34	0.866	0.691	0.536	0.258	0.128
16	2.92	2.58	2.12	1.75	1.34	0.865	0.690	0.535	0.258	0.128
17	2.90	2.57	2.11	1.74	1.33	0.863	0.689	0.534	0.257	0.128
18	2.88	2.55	2.10	1.73	1.33	0.862	0.688	0.534	0.257	0.127
19	2.86	2.54	2.09	1.73	1.33	0.861	0.688	0.533	0.257	0.127
20	2.84	2.53	2.09	1.72	1.32	0.860	0.687	0.533	0.257	0.127
21	2.83	2.52	2.08	1.72	1.32	0.859	0.686	0.532	0.257	0.127
22	2.82	2.51	2.07	1.72	1.32	0.858	0.686	0.532	0.256	0.127
23	2.81	2.50	2.07	1.71	1.32	0.858	0.685	0.532	0.256	0.127
24	2.80	2.49	2.06	1.71	1.32	0.857	0.685	0.531	0.256	0.127
25	2.79	2.48	2.06	1.71	1.32	0.856	0.684	0.531	0.256	0.127
26	2.78	2.48	2.06	1.71	1.32	0.856	0.684	0.531	0.256	0.127
27	2.77	2.47	2.05	1.70	1.31	0.855	0.684	0.531	0.256	0.127
28	2.76	2.47	2.05	1.70	1.31	0.855	0.683	0.530	0.256	0.127
29	2.76	2.46	2.04	1.70	1.31	0.854	0.683	0.530	0.256	0.127
30	2.75	2.46	2.04	1.70	1.31	0.854	0.683	0.530	0.256	0.127
40	2.70	2.42	2.02	1.68	1.30	0.851	0.681	0.529	0.255	0.126
60	2.66	2.39	2.00	1.67	1.30	0.848	0.679	0.527	0.254	0.126
∞	2.62	2.36	1.98	1.66	1.29	0.845	0.677	0.526	0.254	0.126
∞	2.58	2.33	1.96	1.645	1.28	0.842	0.674	0.524	0.253	0.126

Source: R. A. Fisher and F. Yates, *Statistical Tables for Biological, Agricultural and Medical Research* (5th edition), Table III, Oliver and Boyd Ltd., Edinburgh, by permission of the authors and publishers.

Appendix IV

PERCENTILE VALUES (χ_p^2)
for
THE CHI-SQUARE DISTRIBUTION
with v degrees of freedom
(shaded area = p)



v	$\chi_{0.995}^2$	$\chi_{0.99}^2$	$\chi_{0.975}^2$	$\chi_{0.95}^2$	$\chi_{0.90}^2$	$\chi_{0.75}^2$	$\chi_{0.50}^2$	$\chi_{0.25}^2$	$\chi_{0.10}^2$	$\chi_{0.05}^2$	$\chi_{0.025}^2$	$\chi_{0.01}^2$	$\chi_{0.005}^2$
1	7.88	6.63	5.02	3.84	2.71	1.32	0.455	0.102	0.0158	0.0039	0.0010	0.0002	0.0000
2	10.6	9.21	7.38	5.99	4.61	2.77	1.39	0.575	0.211	0.103	0.0506	0.0201	0.0100
3	12.8	11.3	9.35	7.81	6.25	4.11	2.37	1.21	0.584	0.352	0.216	0.115	0.072
4	14.9	13.3	11.1	9.49	7.78	5.39	3.36	1.92	1.06	0.711	0.484	0.297	0.207
5	16.7	15.1	12.8	11.1	9.24	6.63	4.35	2.67	1.61	1.15	0.831	0.554	0.412
6	18.5	16.8	14.4	12.6	10.6	7.84	5.35	3.45	2.20	1.64	1.24	0.872	0.676
7	20.3	18.5	16.0	14.1	12.0	9.04	6.35	4.25	2.83	2.17	1.69	1.24	0.989
8	22.0	20.1	17.5	15.5	13.4	10.2	7.34	5.07	3.49	2.73	2.18	1.65	1.34
9	23.6	21.7	19.0	16.9	14.7	11.4	8.34	5.90	4.17	3.33	2.70	2.09	1.73
10	25.2	23.2	20.5	18.3	16.0	12.5	9.34	6.74	4.87	3.94	3.25	2.56	2.16
11	26.8	24.7	21.9	19.7	17.3	13.7	10.3	7.58	5.58	4.57	3.82	3.05	2.60
12	28.3	26.2	23.3	21.0	18.5	14.8	11.3	8.44	6.30	5.23	4.40	3.57	3.07
13	29.8	27.7	24.7	22.4	19.8	16.0	12.3	9.30	7.04	5.89	5.01	4.11	3.57
14	31.3	29.1	26.1	23.7	21.1	17.1	13.3	10.2	7.79	6.57	5.63	4.66	4.07
15	32.8	30.6	27.5	25.0	22.3	18.2	14.3	11.0	8.55	7.26	6.26	5.23	4.60
16	34.3	32.0	28.8	26.3	23.5	19.4	15.3	11.9	9.31	7.96	6.91	5.81	5.14
17	35.7	33.4	30.2	27.6	24.8	20.5	16.3	12.8	10.1	8.67	7.56	6.41	5.70
18	37.2	34.8	31.5	28.9	26.0	21.6	17.3	13.7	10.9	9.39	8.23	7.01	6.26
19	38.6	36.2	32.9	30.1	27.2	22.7	18.3	14.6	11.7	10.1	8.91	7.63	6.84
20	40.0	37.6	34.2	31.4	28.4	23.8	19.3	15.5	12.4	10.9	9.59	8.26	7.43
21	41.4	38.9	35.5	32.7	29.6	24.9	20.3	16.3	13.2	11.6	10.3	8.90	8.03
22	42.8	40.3	36.8	33.9	30.8	26.0	21.3	17.2	14.0	12.3	11.0	9.54	8.64
23	44.2	41.6	38.1	35.2	32.0	27.1	22.3	18.1	14.8	13.1	11.7	10.2	9.26
24	45.6	43.0	39.4	36.4	33.2	28.2	23.3	19.0	15.7	13.8	12.4	10.9	9.89
25	46.9	44.3	40.6	37.7	34.4	29.3	24.3	19.9	16.5	14.6	13.1	11.5	10.5
26	48.3	45.6	41.9	38.9	35.6	30.4	25.3	20.8	17.3	15.4	13.8	12.2	11.2
27	49.6	47.0	43.2	40.1	36.7	31.5	26.3	21.7	18.1	16.2	14.6	12.9	11.8
28	51.0	48.3	44.5	41.3	37.9	32.6	27.3	22.7	18.9	16.9	15.3	13.6	12.5
29	52.3	49.6	45.7	42.6	39.1	33.7	28.3	23.6	19.8	17.7	16.0	14.3	13.1
30	53.7	50.9	47.0	43.8	40.3	34.8	29.3	24.5	20.6	18.5	16.8	15.0	13.8
40	66.8	63.7	59.3	55.8	51.8	45.6	39.3	33.7	29.1	26.5	24.4	22.2	20.7
50	79.5	76.2	71.4	67.5	63.2	56.3	49.3	42.9	37.7	34.8	32.4	29.7	28.0
60	92.0	88.4	83.3	79.1	74.4	67.0	59.3	52.3	46.5	43.2	40.5	37.5	35.5
70	104.2	100.4	95.0	90.5	85.5	77.6	69.3	61.7	55.3	51.7	48.8	45.4	43.3
80	116.3	112.3	106.6	101.9	96.6	88.1	79.3	71.1	64.3	60.4	57.2	53.5	51.2
90	128.3	124.1	118.1	113.1	107.6	98.6	89.3	80.6	73.3	69.1	65.6	61.8	59.2
100	140.2	135.8	129.6	124.3	118.5	109.1	99.3	90.1	82.4	77.9	74.2	70.1	67.3

Source: Catherine M. Thompson, *Table of percentage points of the χ^2 distribution*,
Biometrika, Vol. 32 (1941), by permission of the author and publisher.