آمار و احتمالات

دکتر هنگامه محمّدی نژاد

گردآورنده : امین دانشی

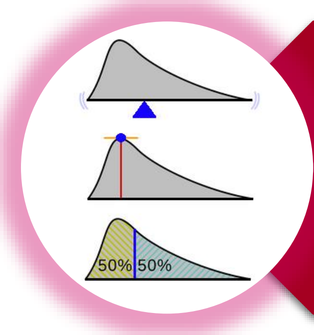
به نام یزدان دانا و توانا

تقدیم به استاد مهربونم، دکتر هنگامه محمّدی نژاد ...

پاییز و زمستان ۹۷

Autumn & Winter 2018

سرفصل‌ها :



فصل اوّل

مشخّصه‌های عددی

- مشخّصه‌های عددی
  - مشخّصه‌های مرکزی
    - میانگین
      - میانگین حسابی
      - میانگین هندسی
      - میانگین هارمونی (همساز)
    - نما (مُد)
    - میانه
  - مشخّصه‌های پراکندگی
    - دامنه تغییرات
    - انحراف چارکی
      - صدک‌ها
      - چارک‌ها
      - دهک‌ها
    - واریانس و انحراف معیار
    - ضریب تغییرات

مشخّصه‌های عددی اعدادی هستند که صفت متغیّر را بصورت خلاصه و اجمال توصیف می‌کنند. این مشخّصه‌ها به دو نوع **مرکزی** و **پراکندگی** وجود دارند.

مشخّصه‌های مرکزی :

اعدادی هستند که مرکز توزیع داده‌ها را نشان می‌دهند و به ۳ نوع تقسیم می‌شوند :

۱. میانگین
۲. نما (مُد)
۳. میانه

۱. میانگین : مهمترین معیار مرکزی است و با توجه به نوع متغیّرها انواع مختلفی دارد.

الف) میانگین حسابی : مهمترین نوع میانگین، میانگین حسابی است که در نمونه با $\bar{x}$ و در جامعه با μ تعریف می‌شود.

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

مثال : میانگین حسابی را محاسبه کنید :

۱. $\mu = \frac{۱۶}{۴} = ۴$ $\qquad x_i = ۲ , ۵ , ۳ , ۶$

۲.

فراوانی

x_i	f_i
۱	۵
۲	۴
۳	۶
	۱۵

$$\mu = \frac{(۵ \times ۱) + (۴ \times ۲) + (۶ \times ۳)}{۱۵} = ۲/۰۶$$

۳.

$$\text{نماینده طبقه} = \frac{\text{کران بالای طبقه + کران پایین طبقه}}{۲}$$

حدود طبقات	f_i	m_i
۰ - ۴	۵	۲
۵ - ۹	۳	۷
۹ - ۱۴	۲	۱۲
۱۵ - ۱۹	۶	۱۷
	۱۶	

$$\mu = \frac{(۵ \times ۲) + (۳ \times ۷) + (۲ \times ۱۲) + (۶ \times ۱۷)}{۱۶} = ۹/۸$$

ب) میانگین هندسی : برای محاسبه متوسّط اندازه‌های نسبی مانند نسبت‌ها، درصدها و شاخص‌ها بکار می‌رود؛ مشروط بر آنکه در محاسبه رشد و نسبت، هر مرحله نسبت به مرحله قبل سنجیده شود و واحد اندازه‌گیری صورت و مخرج نسبت نیز یکی باشد. میانگین هندسی را با μ_G یا $\bar{x}_G$ یا G نشان می‌دهند و بصورت زیر محاسبه می‌شود :

$$G = \sqrt[n]{x_1 \cdot x_2 \ldots x_n}$$

مثال : تولید یک کارخانه طی ۳ سال به ترتیب ۲ ، ۳ و ۵ برابر سال قبل بوده است. متوسّط رشد تولید چقدر است؟

چون بحث نسبت‌هاست پس از

میانگین هندسی استفاده می‌کنیم.

$$G = \sqrt[۳]{۲ \cdot ۳ \cdot ۵} = \sqrt[۳]{۳۰} = ۳/۱۰$$

ج) میانگین هارمونی (همساز) : برای پیدا کردن متوسّط داده‌هایی بکار می‌رود که دارای مقیاس ترکیبی مانند دور بر ساعت، دور بر ثانیه، کیلومتر بر ساعت و بطور کلی صورت و مخرج نسبت‌ها دارای واحدهای اندازه‌گیری متفاوت باشند؛ مشروط بر اینکه مقادیر صورت این نسبت‌ها مساوی باشند. میانگین همساز را با μ_H یا x_H یا H نشان می‌دهند و بصورت زیر محاسبه می‌شود :

$$H = \frac{n}{\sum_{i=۱}^{n} \frac{۱}{x_i}}$$

مثال : ماشین A قطعه‌ای را در عرض ۲ دقیقه و ماشین B در عرض ۴ دقیقه آماده می‌کند. متوسّط تحویل این قطعه چقدر است؟

$$H = \frac{۲}{\frac{۱}{۲} + \frac{۱}{۴}} = \frac{۲}{\frac{۲ + ۱}{۴}} = \frac{۲}{\frac{۳}{۴}} = \frac{۸}{۳}$$

۲. نما (مُد) : متغیّری است که دارای بیشترین تکرار (فراوانی مطلق یا نسبی) است. مُد را با $M.$ نشان می‌دهند و در یک مجموعه از داده‌ها ممکن است مُد وجود نداشته باشد و یک مُد یا بیش از یک مُد وجود داشته باشد. در مجموعه داده‌های پیوسته برای محاسبه مُد از رابطه‌ی زیر استفاده می‌کنند :

طبقه‌ای که دارای بیشترین فراوانی مطلق یا نسبی است طبقه‌ی مُد نام دارد.

فراوانی طبقه مُد

فراوانی طبقه قبل از مُد

$$M. = L_{i} + \frac{f_i - f_{i-1}}{f_i - f_{i-1} + f_i - f_{i+1}} \times L$$

عرض طبقه : *تفاضل کران بالا تا پایین*

فراوانی طبقه بعد از مُد

مثال : مُد را محاسبه کنید :

۱. $M_\circ = ۲$ X : ۱ , ۲ , ۲ , ۵ , ۶ , ۳

$M_\circ =$ نداریم Y : A , B , C , D , E

$M_\circ = A$ و C Z : A , B , C , D , A , C , E , F

$M_\circ =$ نداریم T : A A A , B B B , C C C

۲.

x_i	f_i
۱	۵
۲	۹
۳	۴
۴	۳

چون فراوانی بیشتری دارد.

$M_\circ = ۲$

۳.

حدود طبقات	f_i	کران طبقات
۴ – ۶	۳	۳/۵ – ۶/۵
۷ – ۹	۸	۶/۵ – ۹/۵
۱۰ – ۱۲	۲	۹/۵ – ۱۲/۵
۱۳ – ۱۵	۵	۱۲/۵ – ۱۵/۵

چون فراوانی بیشتری دارد.

$$M_\circ = ۶/۵ + \frac{۸ - ۳}{(۸ - ۳) + (۸ - ۲)} \times ۳$$

$$M_\circ = ۶/۵ + \frac{۵}{۵ + ۶} \times ۳ = ۷/۸۶$$

۳. میانه : اگر داده‌ها بصورت صعودی مرتّب شوند، عددی که در وسط قرار می‌گیرد میانه نام دارد و با M_d نمایش داده می‌شود. اگر تعداد داده‌ها فرد باشد، عدد $\frac{n+1}{2}$ اُمین عدد، میانه را نشان می‌دهد. اگر تعداد داده‌ها زوج باشد، میانگین دو داده‌ی وسط، میانه را نشان می‌دهد.

مثال : در داده‌های زیر میانه را بیابید :

۱.

$X : 2 , 5 , 3 , -1 , 4 , 8 , 1 , 13 , 0$

$X : -1 , 0 , 1 , 2 , 3 , 4 , 5 , 8 , 13$

$M_d = 3$

$M_d = \frac{n+1}{2} = \frac{9+1}{2} = \frac{10}{2} = 5$ پنجمین داده

$Y : 1 , 5 , 9 , 3 , -5 , 2 , 7 , 14$

$Y : -5 , 1 , 2 , 3 , 5 , 7 , 9 , 14$

$M_d = \frac{3+5}{2} = 4$

۲.

فراوانی تجمّعی : با فراوانی نسبی ماقبل خود جمع می‌شود.

x_i	f_i	F_i
۱	۲	۲
۲	۵	۷
۳	۹	۱۶
۴	۴	۲۰
۵	۶	۲۶
	۲۶	

* اوّلین متغیّری که فراوانی تجمّعی آن از $\frac{n+۱}{۲}$ بزرگتر یا مساوی باشد، میانه است.

چون ۱۶ از ۱۳/۵ بزرگتر است، میانه است.

$$\frac{۲۶+۱}{۲} = \frac{۲۷}{۲} = ۱۳/۵$$

$$M_d = ۳$$

۳.

فراوانی طبقه قبل از میانه

$$M_d = L_{mt} + \frac{\frac{n}{۲} - F_c}{f_m} \times L$$

عرض طبقه : تفاضل کران بالا تا پایین

کران پایین طبقه میانه

فراوانی طبقه میانه

* طبقه میانه، اوّلین طبقه‌ای است که فراوانی تجمّعی آن از $\frac{n}{۲}$ بزرگتر و مساوی باشد.

حدود طبقات	f_i	F_i
۴ – ۶	۴	۴
۷ – ۹	۷	۱۱
۱۰ – ۱۲	۳	۱۴
۱۳ – ۱۵	۵	۱۹
	۱۹	

چون ۱۱ از ۹/۵ بزرگتر است پس

این طبقه، طبقه میانه است.

$$\frac{۱۹}{۲} = ۹/۵$$

$$M_d = ۷ + \frac{۹/۵ - ۴}{۷} \times ۲ = ۸/۵$$

مشخصه‌های پراکندگی :

اعدادی هستند که پراکندگی صفت متغیّر را در جامعه اندازه‌گیری می‌کنند و مهمترین آنها عبارتند از :

۱. دامنه تغییرات
۲. انحراف چارکی
۳. واریانس و انحراف معیار
۴. ضریب تغییرات

۱. دامنه تغییرات : اختلاف بزگترین داده و کوچکترین داده است و با R نمایش داده می‌شود.

$$R = \text{Max}\{x_i\} - \text{Min}\{x_i\}$$

مثال : دامنه تغییرات را بیابید :

X : ‎-۱ , ۵ , ۶ , ۲ , ۳

R = ٦ - (-١) = ۷

۲. انحراف چارکی : برای محاسبه‌ی انحراف چارکی لازم است صدک‌ها، دهک‌ها و چارک‌ها را معرّفی کنیم.

صدک‌ها : اعدادی هستند که داده‌ها را به صد قسمت مساوی تقسیم می‌کنند و اگر P عددی بین صفر تا صد باشد، صدک P اُم را با H_P نمایش می‌دهند و عددی است که P٪ داده‌ها از آن کمتر و P٪ - ۱ داده‌ها از آن بیشتر است. برای بدست آوردن H_P، داده‌ها را بصورت صعودی مرتّب می‌کنیم، عدد $\frac{(n+۱)P}{۱۰۰}$ را محاسبه می کنیم؛ اگر این عدد برابر با یک عدد صحیح مثل r باشد، آنگاه r اُمین داده برابر H_P است و اگر یک عدد غیر صحیح باشد، قسمت صحیح آن را r و قسمت اعشار آن را w درنظر می‌گیریم، آنگاه H_P از رابطه‌ی زیر بدست می‌آید :

$$H_P = (۱ - w) . x_r + w . x_{r+۱}$$

که در آن x_r ، r اُمین داده مرتّب می‌باشد.

مثال : صدک ۴۳ اُم را بدست آورید :

X : ۲ , ۴ , ۱ , ۴ , ۹ , -۱ , ۳ , ۷ , ۱۶ , ۶ , ۵ ,۱۲

X : -۱ , ۱ , ۲ , ۳ , ۴ , ۴ , ۵ , ۶ , ۷ , ۹ , ۱۲ , ۱۶

$$H_{۴۳} = \frac{(n+۱)P}{۱۰۰} = \frac{(۱۲+۱)۴۳}{۱۰۰} = \frac{۱۳ \times ۴۳}{۱۰۰} = \underset{r}{۵}/\underbrace{۵۹}_{w}$$

$$H_{۴۳} = (۱ - w) . x_r + w . x_{r+۱}$$

$$H_{۴۳} = (۱ - ۰/۵۹) \times x_۵ + ۰/۵۹ \times x_۶$$

$$H_{۴۳} = ۰/۴۱ \times ۴ + ۰/۵۹ \times ۴ = ۴$$

چارک‌ها : اعدادی هستند که چنانچه داده‌ها را بصورت صعودی مرتّب کنیم و داده‌ها را به ۴ قسمت مساوی تقسیم کنیم، آنگاه چارک اوّل (Q_1) اعدادی هستند که $\frac{۱}{۴}$ داده‌ها از آن‌ها کمتر و $\frac{۳}{۴}$ از آن‌ها بیشتر است. واضح است که $Q_1 = H_{25}$، چارک دوّم (Q_2) اعدادی هستند که $\frac{۲}{۴}$داده‌ها از آن‌ها کمتر و $\frac{۲}{۴}$ از آن‌ها بیشتر است. واضح است که $Q_2 = H_{50} = M_d$ و چارک سوّم (Q_3) اعدادی هستند که $\frac{۳}{۴}$داده‌ها از آن‌ها کمتر و $\frac{۱}{۴}$ از آن‌ها بیشتر است. واضح است که $Q_3 = H_{75}$.

مثال : در داده‌های زیر چارک اوّل و سوّم را محاسبه کنید.

X : ۲ , ۵ , -۱ , ۴ , ۸ , ۱۳ , ۹ , ۱۷

X : -۱ , ۲ , ۴ , ۵ , ۸ , ۹ , ۱۳ , ۱۷

$$Q_1 = \frac{(n+1)P}{100} = \frac{(8+1)25}{100} = \frac{9 \times 25}{100} = 2/25$$

$$Q_1 = (1 - W) . x_r + W . x_{r+1}$$

$$Q_1 = (1 - 0/25) \times x_2 + 0/25 \times x_3$$

$$Q_1 = 0/75 \times 2 + 0/25 \times 4 = 2/5$$

$$Q_3 = \frac{(n+1)P}{100} = \frac{(8+1)75}{100} = \frac{9 \times 75}{100} = 6/25$$

$$Q_3 = (1 - W) . x_r + W . x_{r+1}$$

$$Q_3 = (1 - 0/75) \times x_6 + 0/75 \times x_7$$

$$Q_3 = 0/25 \times 9 + 0/75 \times 13 = 12$$

دهک‌ها : اعدادی هستند که چنانچه داده‌ها را بصورت صعودی مرتّب کنیم و داده‌ها را به ۱۰ قسمت مساوی تقسیم کنیم، آنگاه دهک اوّل ($D_{۱}$) اعدادی هستند که $\frac{۱}{۱۰}$ داده‌ها از آن‌ها کمتر و $\frac{۹}{۱۰}$ از آن‌ها بیشتر است. واضح است که $D_{۱} = H_{۱۰}$ و دهک پنجم ($D_{۵}$) اعدادی هستند که $\frac{۵}{۱۰}$ داده‌ها از آن‌ها کمتر و $\frac{۵}{۱۰}$ از آن‌ها بیشتر است. واضح است که $D_{۵} = H_{۵۰} = M_d$.

مثال : در داده‌های زیر دهک هفتم را محاسبه کنید.

X : ۲ , ۵ , ۱ , ۰ , ۶ , ۹ , -۱ , ۴ , ۸

X : -۱ , ۰ , ۱ , ۲ , ۴ , ۵ , ۶ , ۸ , ۹

$$D_{۷} = \frac{(n+۱)P}{۱۰۰} = \frac{(۹+۱)۷۰}{۱۰۰} = \frac{۱۰ \times ۷۰}{۱۰۰} = ۷$$

$$D_{۷} = ۶$$

مثال : صدک چهل و پنجم، چارک سوّم و دهک سوّم را محاسبه کنید.

x_i	f_i	F_i
۰	۲	۲
۱	۵	۷
۲	۷	۱۴
۳	۹	۲۳
۴	۳	۲۶
۵	۸	۳۴
۶	۴	۳۸
۷	۱۰	۴۸
	۴۸	

$$D_{۳} = \frac{(n+۱)P}{۱۰۰} = \frac{(۴۸+۱)\,۳۰}{۱۰۰} = \frac{۴۹ \times ۳۰}{۱۰۰} = ۱۴/۷$$

$$D_{۳} = (۱ - W) . x_{r} + W . x_{r+۱}$$

$$D_{۳} = (۱ - ۰/۷) \times x_{۱۴} + ۰/۷ \times x_{۱۵}$$

$$D_{۳} = ۰/۳ \times ۲ + ۰/۷ \times ۳ = ۲/۷$$

$$Q_{۳} = \frac{(n+۱)P}{۱۰۰} = \frac{(۴۸+۱)\,۷۵}{۱۰۰} = \frac{۴۹ \times ۷۵}{۱۰۰} = ۳۶/۷۵$$

$$Q_{۳} = ۶$$

$$H_{۴۵} = \frac{(n+۱)P}{۱۰۰} = \frac{(۴۸+۱)\,۴۵}{۱۰۰} = \frac{۴۹ \times ۴۵}{۱۰۰} = ۲۲/۵$$

$$H_{۴۵} = ۳$$

---

۳. واریانس و انحراف معیار : مهم‌ترین معیارهای پراکندگی در آمار، واریانس و انحراف معیار می‌باشند. واریانس را در جامعه با σ^2 و در نمونه با S^2 نمایش می‌دهند و بصورت زیر محاسبه می‌شود :

$$\sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}$$

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

* واریانس نمونه‌ای را بصورت زیر تعریف می‌کنند :

$$S^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}}{n-1}$$

* جذر مثبت واریانس را انحراف معیار می‌گویند که در جامعه با σ و در نمونه با S نمایش می‌دهند.

* اگر داده‌های گسسته دارای فراوانی باشد، برای محاسبه‌ی واریانس از رابطه‌ی زیر استفاده می‌شود :

$$S^2 = \frac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}{n-1}$$

یا

$$S^2 = \frac{\sum_{i=1}^{n} f_i . x_i^2 - \frac{(\sum_{i=1}^{n} f_i . x_i)^2}{n}}{n-1}$$

* در داده‌های پیوسته‌ی طبقه‌بندی شده‌ی با نماینده‌ی طبقات m_i و فراوانی f_i از رابطه‌ی زیر برای محاسبه‌ی واریانس استفاده می‌شود :

$$S^{۲} = \frac{\sum_{i=۱}^{n} f_i (m_i - \bar{x})^{۲}}{n-۱}$$

یا

$$S^{۲} = \frac{\sum_{i=۱}^{n} f_i . m_i^{۲} - \frac{(\sum_{i=۱}^{n} f_i . m_i)^{۲}}{n}}{n-۱}$$

مثال : در هر یک از داده‌های زیر واریانس و انحراف معیار را بدست آورید :

۱. $X : ۲ , ۵ , ۱ , ۴$

$$S^{۲} = \frac{\sum_{i=۱}^{n} x_i^{۲} - \frac{(\sum_{i=۱}^{n} x_i)^{۲}}{n}}{n-۱}$$

$$S^{۲} = \frac{۴+۲۵+۱+۱۶ - \frac{(۲+۵+۱+۴)^{۲}}{۴}}{۴-۱}$$

$$S^{۲} = \frac{۴۶ - \frac{۱۴۴}{۴}}{۳} = ۳/۳$$

۲.

x_i	f_i
۱	۵
۲	۴
۳	۱
	۱۰

$$S^{۲} = \frac{\sum_{i=۱}^{n} f_i . x_i^{۲} - \frac{\left(\sum_{i=۱}^{n} f_i . x_i\right)^{۲}}{n}}{n-۱}$$

$$S^{۲} = \frac{(۵ \times ۱^{۲}) + (۴ \times ۲^{۲}) + (۱ \times ۳^{۲}) - \frac{((۵ \times ۱) + (۴ \times ۲) + (۱ \times ۳))^{۲}}{۱۰}}{۱۰ - ۱}$$

$$S^{۲} = \frac{(۵ + ۱۶ + ۹) - \frac{(۵ + ۸ + ۳)^{۲}}{۱۰}}{۹}$$

$$S^{۲} = \frac{۳۰ - \frac{۲۵۶}{۱۰}}{۹} = ۰/۴۸$$

۳.

حدود طبقات	f_i	m_i
۲ - ۴	۴	۳
۵ - ۷	۳	۶
۸ - ۱۰	۳	۹
	۱۰	

$$S^۲ = \frac{\sum_{i=۱}^{n} f_i . m_i^۲ - \frac{(\sum_{i=۱}^{n} f_i . m_i)^۲}{n}}{n - ۱}$$

$$S^۲ = \frac{(۴ \times ۳^۲) + (۳ \times ۶^۲) + (۳ \times ۹^۲) - \frac{((۴ \times ۳) + (۳ \times ۶) + (۳ \times ۹))^۲}{۱۰}}{۱۰ - ۱}$$

$$S^۲ = \frac{(۳۶ + ۱۰۸ + ۲۴۳) - \frac{(۱۲ + ۱۸ + ۲۷)^۲}{۱۰}}{۹}$$

$$S^۲ = \frac{۳۸۷ - \frac{۳۲۴۹}{۱۰}}{۹} = ۶/۹$$

فصل دوّم

توزیع احتمال

- توزیع احتمال
  - توزیع احتمال گسسته
    - توزیع احتمال برنولی
    - توزیع دو جمله‌ای
    - توزیع احتمال پوآسون
    - توزیع احتمال فوق هندسی
  - توزیع احتمال پیوسته
    - توزیع یکنواخت
    - توزیع نمایی
    - توزیع نرمال
    - توزیع نرمال استاندارد

توزیع احتمال گسسته :

توزیع احتمال برنولی :

آزمایش تصادفی برنولی آزمایشی است که در آن فقط ۲ نتیجه اتّفاق می‌افتد؛ پیروزی و شکست. یک نتیجه را به دلخواه پیروزی و دیگری را شکست در نظر می‌گیریم. احتمال پیروزی را با p و احتمال شکست را با $q = ۱ - p$ نمایش می‌دهیم و تابع احتمال توزیع برنولی را بصورت زیر تعریف می‌کنیم :

$$P (X = x) = p^{x} (۱ - p)^{۱ - x} \qquad x = ۰ , ۱$$

امید ریاضی توزیع برنولی را $E (X) = p$ و واریانس آن را $V (X) = p (۱ - p)$ تعریف می‌کنند.

توزیع دو جمله‌ای :

اگر آزمایش برنولی را n بار مستقل از هم تکرار کنیم و احتمال پیروزی (p) در طول آزمایش ثابت باشد، به آن آزمایش دو جمله‌ای می‌گویند. متغیّر تصادفی X که نشان دهنده‌ی تعداد پیروزی‌ها در این n آزمایش است را متغیّر تصادفی دو جمله‌ای می‌گویند که دارای تابع توزیع احتمال زیر است :

$$P (X) = C^{n}_{x} \; p^{x} (۱ - p)^{n - x} \qquad x = ۰ , ۱ , ... , n$$

مثال : احتمال رو آمدن در یک سکّه $\frac{۱}{۳}$ است. سکّه‌ای را ۵ بار پرتاب می‌کنیم؛

الف) احتمال اینکه ۲ بار رو بیاید چقدر است؟

ب) احتمال اینکه ۳ بار پشت بیاید چقدر است؟

ج احتمال اینکه حدّاقل ۱ بار رو بیاید چقدر است؟

$$p_H=\frac{۱}{۳}\quad,\quad p_T=\frac{۲}{۳}\quad,\quad n=۵$$

الف)

$$P(X=۲)=C_{۲}^{۵}\left(\frac{۱}{۳}\right)^{۲}\left(\frac{۲}{۳}\right)^{۳}$$

$$\binom{۵}{۲}\left(\frac{۱}{۳}\right)^{۲}\left(\frac{۲}{۳}\right)^{۳}$$

$$\frac{۵!}{۲!\,(۵-۲)!}\times\frac{۸}{۳^{۵}}$$

$$\frac{۵\times۴\times\cancel{۳!}}{۲\times۱!\times\cancel{۳!}}\times\frac{۸}{۳^{۵}}=\frac{۸۰}{۳^{۵}}$$

ب)

$$P(X=۳)=C_{۳}^{۵}\left(\frac{۲}{۳}\right)^{۳}\left(\frac{۱}{۳}\right)^{۲}$$

$$\binom{۵}{۳}\left(\frac{۲}{۳}\right)^{۳}\left(\frac{۱}{۳}\right)^{۲}$$

$$\frac{۵!}{۳!\,(۵-۳)!}\times\frac{۸}{۳^{۵}}$$

$$\frac{۵\times۴\times\cancel{۳!}}{\cancel{۳!}\times۲\times۱!}\times\frac{۸}{۳^{۵}}=\frac{۸۰}{۳^{۵}}$$

ج)

$$P(X \geq 1) = 1 - p(X = 0)$$

$$1 - \binom{5}{0} \cancel{\left(\frac{1}{3}\right)^0} \left(\frac{2}{3}\right)^5$$

$$1 - \frac{\cancel{5!}}{\underset{1}{\cancel{0! \times 5!}}} \times \frac{2^5}{3^5} = \left(\frac{2}{3}\right)^5$$

مثال : احتمال اینکه تیراندازی تیر خود را به هدف بزند $\frac{3}{4}$ است. احتمال اینکه از ۴ تیری که شلّیک خواهد کرد دقیقاً ۲ تیر به هدف بخورد چقدر است؟

$$p_H = \frac{3}{4} \quad , \quad p_T = q = 1 - p = \frac{1}{4} \quad , \quad n = 4$$

$$P(X = 2) = C_2^4 \left(\frac{3}{4}\right)^2 \left(\frac{1}{4}\right)^2$$

$$\binom{4}{2} \left(\frac{3}{4}\right)^2 \left(\frac{1}{4}\right)^2$$

$$\frac{4!}{2!\,(4-2)!} \times \frac{9}{16} \times \frac{1}{16}$$

$$\frac{4 \times 3 \times \cancel{2!}}{2 \times 1! \times \cancel{2!}} \times \frac{9}{16 \times 16} = \frac{54}{256}$$

مثال : خانواده‌ای دارای ۶ فرزند است. مطلوب است محاسبه احتمال اینکه :

الف) حدّاکثر ۲ فرزند پسر وجود داشته باشد.

ب) حدّاقل ۲ فرزند پسر وجود داشته باشد.

ج) دقیقاً ۲ فرزند پسر وجود داشته باشد.

$$p_H = \frac{1}{2} \quad , \quad p_T = q = 1 - p = \frac{1}{2} \quad , \quad n = 6$$

الف)

$$P(X \le 2) = \sum_{x=0}^{6} p(X) = p(X=0) + p(X=1) + p(X=2)$$

$$\binom{6}{0}\left(\frac{1}{2}\right)^{0}\left(\frac{1}{2}\right)^{6} + \binom{6}{1}\left(\frac{1}{2}\right)^{1}\left(\frac{1}{2}\right)^{5} + \binom{6}{2}\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{4}$$

ب)

$$P(X \ge 2) = \sum_{x=2}^{6} p(X) = p(X=2) + p(X=3) + p(X=4) + p(X=5) + p(X=6)$$

$$\binom{6}{2}\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{4} + \binom{6}{3}\left(\frac{1}{2}\right)^{3}\left(\frac{1}{2}\right)^{3}$$

$$+\binom{6}{4}\left(\frac{1}{2}\right)^{4}\left(\frac{1}{2}\right)^{2} + \binom{6}{5}\left(\frac{1}{2}\right)^{5}\left(\frac{1}{2}\right)^{1} + \binom{6}{6}\left(\frac{1}{2}\right)^{6}\left(\frac{1}{2}\right)^{0}$$

ج)

$$P(X=2) = C_{2}^{6}\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{4}$$

$$\binom{6}{2}\left(\frac{1}{2}\right)^{2}\left(\frac{1}{2}\right)^{4}$$

مثال : فرض کنید که ۴۰٪ افراد شاغل خواستار ثابت ماندن دستمزد و کنترل قیمت‌ها هستند. اگر ۵ نفر از این افراد را به تصادف انتخاب کنیم، احتمال اینکه :

الف) حدّاقل ۳ نفر خواستار ثابت ماندن دستمزدها باشند چقدر است؟

ب) احتمال اینکه هیچکدام موافق با این قضیه نباشند چقدر است؟

$$p_H = \frac{4}{10} \quad , \quad p_T = q = 1 - p = \frac{6}{10} \quad , \quad n = 5$$

الف)

$$P(X \geq 3) = \sum_{x=3}^{5} p(X) = p(X=3) + p(X=4) + p(X=5)$$

$$\binom{5}{3}\left(\frac{4}{10}\right)^3 \left(\frac{6}{10}\right)^2 + \binom{5}{4}\left(\frac{4}{10}\right)^4 \left(\frac{6}{10}\right)^1 + \binom{5}{5}\left(\frac{4}{10}\right)^5 \left(\frac{6}{10}\right)^0$$

ب)

$$P(X = 0) = C_0^5 \left(\frac{4}{10}\right)^0 \left(\frac{6}{10}\right)^0$$

$$\binom{5}{0}\left(\frac{4}{10}\right)^0 \left(\frac{6}{10}\right)^0$$

میانگین، واریانس و انحراف معیار متغیّر تصادفی دو جمله‌ای :

میانگین، واریانس و انحراف معیار متغیّر تصادفی دو جمله‌ای X با پارامترهای n و p بصورت زیر می‌باشد :

میانگین X $\mu = E(X) = n \cdot p$

واریانس X $\sigma^{2} = V(X) = n \cdot p \cdot q$

انحراف معیار X $\sigma = \sqrt{n \cdot p \cdot q}$

مثال : گزارشی شده است که در شهری ۷۵٪ خانه‌ها با گاز شهری گرم می‌شوند. اگر در یک محلّه ۲۵۰۰ خانه وجود داشته باشد، انتظار داریم تعداد خانه‌هایی که بوسیله‌ی گاز شهری گرم می‌شوند چقدر باشد؟

$\mu = E(X) = n \cdot p$

۲۵۰۰ × ٪۷۵ = ۱۸۷۵

اگر X تعداد خانه‌های این محلّه باشد که با گاز گرم می‌شوند، واریانس و انحراف معیار X را بدست آورید :

$\sigma^{2} = V(X) = n \cdot p \cdot q$

۲۵۰۰ × ٪۷۵ × ٪۲۵ = ۴۶۸/۷

$\sigma = \sqrt{n \cdot p \cdot q}$ = √۴۶۸/۷

توزیع احتمال پوآسون :

این توزیع احتمال مدل خوبی برای داده‌هایی است که معرّف تعداد وقوع پیشامدی معیّن در یک زمان واحد یا مکان واحد باشد. در این نوع توزیع X تعداد پیشامدهایی است که در یک دوره زمانی یا مکانی رخ دهد و پیشامدها بطور تصادفی و مستقل از یکدیگر رخ می‌دهند. در توزیع پوآسون " μ " میانگین توزیع احتمال پوآسون است. بنابراین اگر تعداد متوسّط متغیّر تصادفی X که توزیع پوآسون دارد معلوم باشد، توزیع احتمال X و واریانس X نیز بدست خواهد آمد. فرمول توزیع احتمال پوآسون بصورت زیر است :

$$P(X = x) = \frac{\mu^{x} e^{-\mu}}{x!} \qquad x = ۰ , ۱ , ۲ , ...$$

میانگین $\mu_x = \mu$

واریانس $\sigma_x^۲ = \mu$

* در مواردی که بجای توزیع دو جمله‌ای از توزیع پوآسون استفاده می‌شود، می‌توان نکات زیر را در نظر گرفت :

۱. در آزمایش‌های دو جمله‌ای که $n > ۲۰$ باشد.

۲. p بقدر کافی کوچک باشد. ($np \leq ۵$)

در این شرایط متغیّر تصادفی دو جمله‌ای دارای توزیع پوآسون با مقدار متوسّط $\mu = np$ می‌باشد.

مثال : اگر تعداد مراجعین به بانکی بطور متوسّط ۷۲ نفر در ساعت باشد،

الف) احتمال اینکه ۴ نفر در ۱۰ دقیقه اوّل ساعتی به بانک مراجعه کنند چقدر است؟

ب) احتمال اینکه در ۵ دقیقه آخر ساعت هیچکس به بانک مراجعه نکند چقدر است؟

الف)

۶۰	۷۲
۱۰	؟

$$\mu = \frac{72 \times 10}{60} = 12$$

$$P(X = 4) = \frac{12^{4} e^{-12}}{4!}$$

ب)

۶۰	۷۲
۵	؟

$$\mu = \frac{72 \times 5}{60} = 6$$

$$P(X = 0) = \frac{6^{0} e^{-6}}{0!}$$

مثال : ۲۰۰ مسافر برای هواپیما جا رزرو کرده‌اند. اگر احتمال مراجعه نکردن مسافر $\frac{1}{100}$ باشد،

الف) احتمال اینکه حدّاقل ۱ نفر مراجعه نکند چقدر است؟

ب) احتمال اینکه ۳ نفر مراجعه نکنند چقدر است؟

$$p = \frac{1}{100} \quad , \quad n = 200$$

الف)

$$\mu = 200 \times \frac{1}{100} = 2$$

$$P(X \geq 1) = 1 - p(X = 0)$$

$$1 - \frac{2^{0} e^{-2}}{0!}$$

ب)

$$P(X = 3) = \frac{2^{3} e^{-2}}{3!}$$

مثال : اگر تعداد افرادی که به بخش معیّنی از یک بیمارستان در یک روز مراجعه می‌کنند، متغیّر تصادفی **x** از نوع پوآسون باشد،

الف) احتمال اینکه تعداد افرادی که به این بخش مراجعه کنند در یک روز معیّن ۲ نفر باشند، چقدر است؟

ب) احتمال اینکه **x** حدّاکثر ۲ باشد چقدر است؟

$$\mu = ۵$$

الف)

$$P(X=۲)=\frac{۵^{۲}e^{-۵}}{۲!}$$

ب)

$$P(X\le ۲)=p(X=۰)+p(X=۱)+p(X=۲)$$

$$\frac{۵^{۰}e^{-۵}}{۰!}+\frac{۵^{۱}e^{-۵}}{۱!}+\frac{۵^{۲}e^{-۵}}{۲!}$$

توزیع احتمال فوق هندسی :

اگر بخواهیم یک نمونه از جامعه‌ای را انتخاب کنیم و ببینیم که چه تعدادی از اعضای نمونه دارای مشخّصه‌ی معیّنی هستند، اگر انتخاب نمونه بدون جایگذاری انجام شود ولی تعداد عضوهای نمونه نسبت به عضوهای جامعه خیلی کم باشد ($n \leq ۵٪ \, N$)، این انتخاب یک آزمایش دو جمله‌ای است و فقط در مواردی که انتخاب نمونه بدون جایگذاری انجام شود و اندازه‌ی نمونه نسبت به اندازه‌ی جامعه بزرگ باشد ($n > ۵٪ \, N$)، از توزیع فوق هندسی استفاده می‌کنیم. در این حالت موفّقیّت در یک آزمایش به نتیجه‌ی آزمایش‌های قبل بستگی دارد و x تعداد موفّقیّت‌ها در نمونه‌ی یک متغیّر تصادفی فوق هندسی می‌باشد که توزیع احتمال آن بصورت زیر تعریف می‌شود :

$$P(x) = \frac{C_x^k \, C_{n-x}^{N-k}}{C_n^N} \qquad x = ۰ , ۱ , ۲ , ... , n$$

n : تعداد نمونه

N : تعداد جامعه

k : تعداد اعضای جامعه دارای صفت مشخّصه (موفّقیّت)

x: تعداد موفّقیّت‌ها در نمونه

میانگین x

$$\mu_x = E(x) = n \frac{k}{N}$$

واریانس x

$$\sigma_x^۲ = V(x) = n \left(\frac{k}{N}\right) \left(\frac{N-k}{N}\right) \left(\frac{N-n}{N-۱}\right)$$

مثال : فرض کنید **x** تعداد موفّقیّت‌های مشاهده در یک نمونه‌ی ۵ تایی انتخاب شده از یک جامعه‌ی ۱۰ عضوی باشد. اگر در این جامعه ۶ عضو دارای مشخّصه‌ی موفّقیّت باشد،

الف) احتمال اینکه هیچ موفّقیّتی مشاهده نشود چقدر است؟

ب) احتمال اینکه حدّاقل ۲ موفّقیّت مشاهده شود چقدر است؟

ج) احتمال اینکه دقیقاً ۲ موفّقیّت مشاهده شود چقدر است؟

$N : 10$

$n : 5$

$k : 6$

الف)

$$P(X=0)=0$$

ب)

$$P(X \geq 2) = 1 - p(X < 2)$$

$$1 - \left(\cancel{p(X=0)} + p(X=1) \right)$$

$$1 - p(X=1)$$

$$1 - \frac{C^{6}_{1}\, C^{10-6}_{5-1}}{C^{10}_{5}}$$

$$1 - \frac{\binom{6}{1}\binom{4}{4}}{\binom{10}{5}}$$

ج)

$$P(X=2) = \frac{C^{6}_{2}\, C^{10-6}_{5-2}}{C^{10}_{5}}$$

$$\frac{\binom{6}{2}\binom{4}{3}}{\binom{10}{5}}$$

توزیع احتمال پیوسته :

توزیع یکنواخت :

این توزیع مدل خوبی برای توزیع احتمال متغیّری است که بطور تصادفی هر یک از مقادیر بین دو نقطه‌ی (b و a) که (a < b) را روی خط اعداد حقیقی اختیار می‌کند. تابع چگالی احتمال یکنواخت با متغیّر تصادفی یکنواخت x برابر است با :

$$\begin{cases} \dfrac{۱}{b - a} & a < x < b \\ ۰ & \text{سایر نقاط} \end{cases}$$

و میانگین آن عبارت است از :

$$\mu = \frac{a + b}{۲}$$

* اگر مقادیر c و d طوری باشند که a < c < d < b آنگاه :

$$P(c \leq x \leq d) = \int_c^d \frac{۱}{b - a}\, dx$$

$$\frac{۱}{b - a}\, x \Big|_c^d$$

$$\frac{۱}{b - a}\,(d - c)$$

باید یکبار c و یکبار d را بجای x جایگذاری می‌کردیم که بجای آن فاکتور گرفتیم 😎

یادآوری انتگرال : $\int_a^b f(x)dx = F(b) - F(a)$

مثال : متغیّر تصادفی یکنواخت x در بازه بین ۲/۵ تا ۴/۵ تعریف شده است.

الف) تابع احتمال این متغیّر تصادفی را بدست آورید.

ب) میانگین این متغیّر تصادفی را بدست آورید.

ج) احتمال اینکه ۴ > x > ۳ باشد را بدست آورید.

د) احتمال اینکه ۶ > x > ۴ باشد را بدست آورید.

الف)

$$f(x) = \begin{cases} \dfrac{۱}{۴/۵ - ۲/۵} = \dfrac{۱}{۲} & ۲/۵ < x < ۴/۵ \\ ۰ & \text{سایر نقاط} \end{cases}$$

ب)

$$\mu = \frac{a+b}{۲} = \frac{۲/۵ + ۴/۵}{۲} = \frac{۷}{۲} = ۳/۵$$

ج)

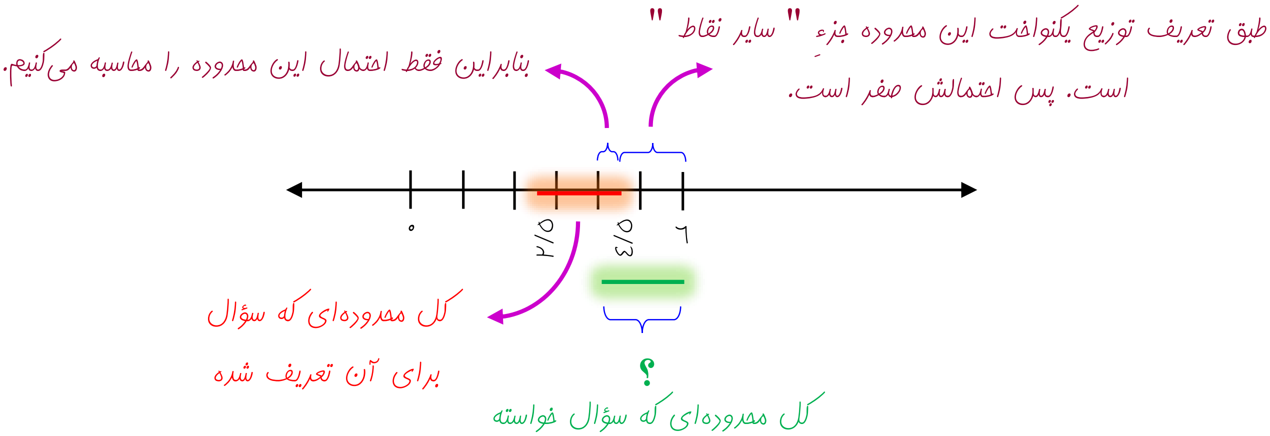
$$P(۳ < x < ۴) = \int_{۳}^{۴} \frac{۱}{۲} dx$$

چون احتمال در محدوده‌ی تعریف شده است، پس از همان $\frac{۱}{۲}$ استفاده می‌کنیم.

$$\frac{۱}{۲} x \Big|_{۳}^{۴}$$

$$\frac{۱}{۲}(۴ - ۳) = \frac{۱}{۲}$$

د)



$$P(٤ < x < ٦) = P(٤ < x < ٤/٥) + P(٤/٥ < x < ٦)$$

$$\int_{٤}^{٤/٥} \frac{١}{٢}\, dx + \int_{٤/٥}^{٦} ٠\, dx$$

$$\frac{١}{٢}\, x \Big|_{٤}^{٤/٥}$$

$$\frac{١}{٢}\,(٤/٥ - ٤)$$

$$\frac{١}{٢}\left(\frac{١}{٢}\right) = \frac{١}{٤}$$

توزیع نمایی :

متغیّر تصادفی نمایی x که دارای چگالی احتمال زیر می‌باشد را یک متغیّر نمایی با پارامتر توزیع نمایی λ در نظر می‌گیریم :

$$\begin{cases} \lambda e^{-\lambda x} & 0 \leq x \quad , \quad 0 < \lambda \\ 0 & \text{سایر نقاط} \end{cases}$$

میانگین و واریانس تابع نمایی برابر است با :

$$\mu = \frac{1}{\lambda}$$

$$\sigma^{2} = \frac{1}{\lambda^{2}}$$

* وقتی تعداد وقوع حوادث در یک مدّت زمانیِ مشخّص توزیع پوآسون با میانگین λ دارد، زمان انتظار بین وقوع دو حادثه متوالی توزیع نمایی با پارامتر λ دارد.

مثال : متغیّر تصادفی x دارای توزیع نمایی با تابع چگالی احتمال زیر است :

$$\begin{cases} \frac{۱}{۴} e^{-\frac{۱}{۴}x} & ۰ \leq x \\ ۰ & \text{سایر نقاط} \end{cases}$$

الف) $P(x = ۴)$

ب) $P(x \leq ۱)$

ج) $P(x \geq ۱/۵)$

د) $P(-۱/۵ \leq x \leq ۳)$

الف)

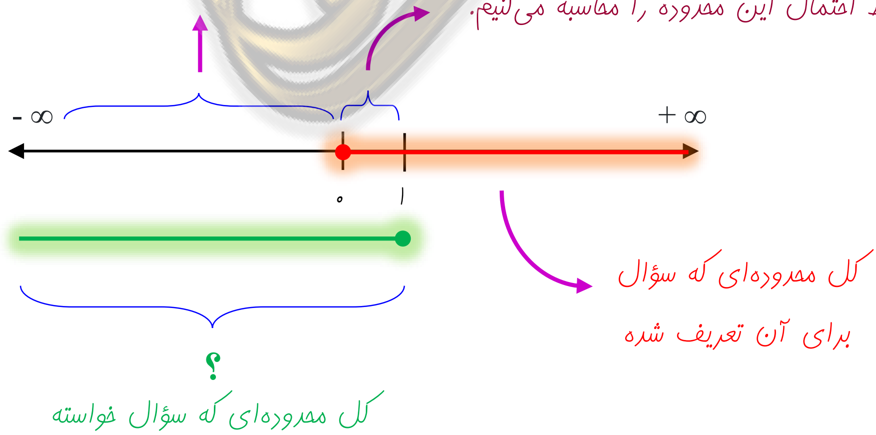
انتگرال از خودِ یک عدد به خودش برابر است با صفر

$$P(x = ۴) = \int_{۴}^{۴} \frac{۱}{۴} e^{-\frac{۱}{۴}x}\, dx = ۰$$

ب)

طبق تعریف توزیع نمایی این محدوده جزء " سایر نقاط " است. پس احتمالش صفر است.

بنابراین فقط احتمال این محدوده را محاسبه می‌کنیم.



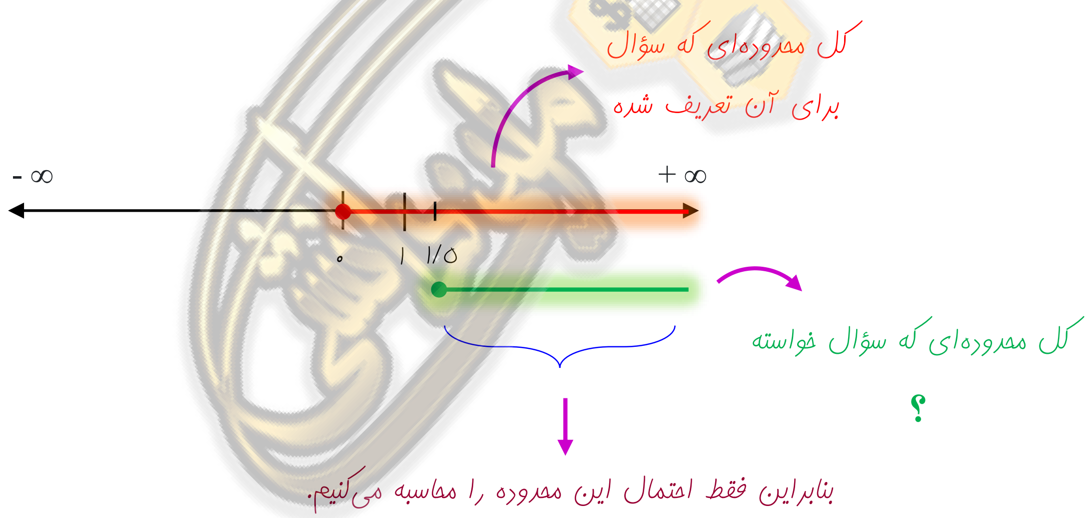
$$P(x \le 1) = P(\cancel{x < 0})^{0} + P(0 \le x \le 1)$$

$$\int_0^1 \frac{1}{4} e^{-\frac{1}{4}x}\, dx$$

$$-e^{-\frac{1}{4}x}\Big|_0^1$$

$$-\left(\underbrace{\cancel{e^{-\frac{1}{4}(1)}}}_{e^{-\frac{1}{4}}} - \overbrace{\cancel{e^{-\frac{1}{4}(0)}}}^{1}\right) = 1 - e^{-\frac{1}{4}}$$

ج)

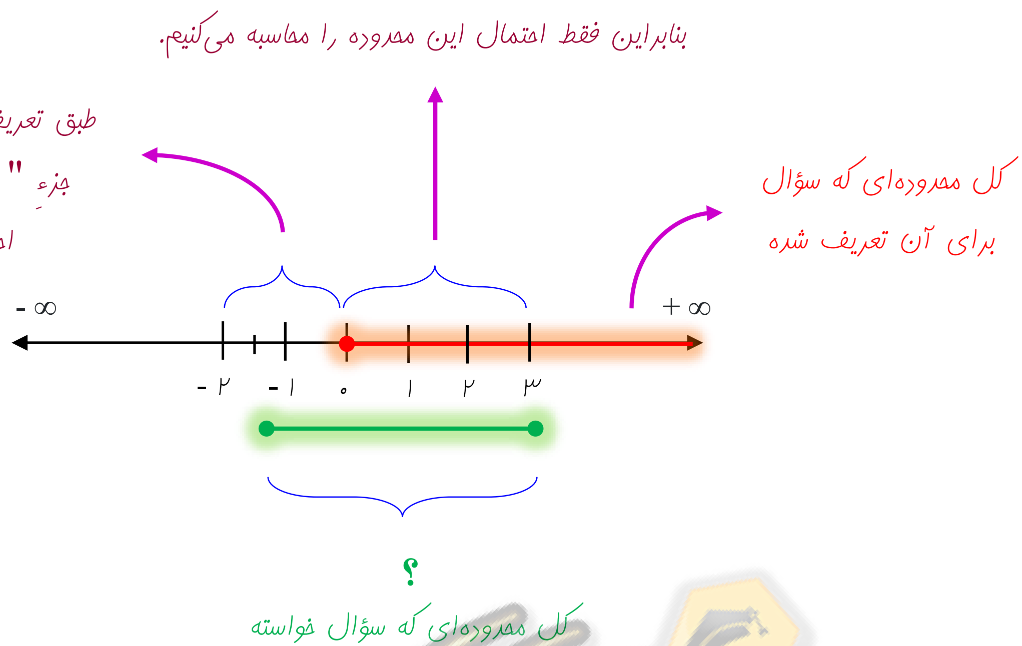


$$P(x \ge 1.5) = \int_{1.5}^{+\infty} \frac{1}{4} e^{-\frac{1}{4}x}\, dx$$

$$-e^{-\frac{1}{4}x}\Big|_{1.5}^{+\infty}$$

$$-\left(\underbrace{\cancel{e^{-\frac{1}{4}(+\infty)}}}_{\cancel{e^{-\infty}} \to 0} - e^{-\frac{1}{4}\left(\frac{3}{2}\right)}\right) = e^{-\frac{3}{8}}$$

د)



$$P(-1/5 \le x \le 3) = \cancel{P(-1/5 \le x < 0)}^{\,0} + P(0 \le x \le 3)$$

$$\int_0^3 \frac{1}{4} e^{-\frac{1}{4}x}\,dx$$

$$-e^{-\frac{1}{4}x}\Big|_0^3$$

$$-\left(\cancel{e^{-\frac{1}{4}(3)}} - \cancel{e^{-\frac{1}{4}(0)}}^{\,1}\right) = 1 - e^{-\frac{3}{4}}$$

$$e^{-\frac{3}{4}}$$

مثال : طول عمر قطعه‌ای از یک ماشین دارای توزیع نمایی با میانگین ۷،۵۰۰ ساعت است.

الف) احتمال اینکه این قطعه ۵،۰۰۰ ساعت عمر کند چقدر است؟

ب) احتمال اینکه این قطعه بیش از ۱۰،۰۰۰ ساعت عمر کند چقدر است؟

ج) احتمال اینکه این قطعه کمتر از ۲،۵۰۰ ساعت عمر کند چقدر است؟

$$\mu = \frac{1}{\lambda}$$

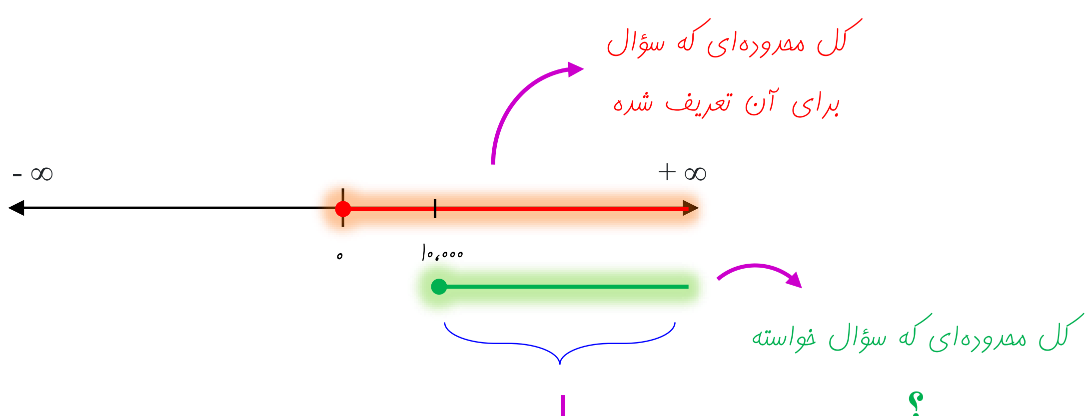
$$7{,}500 = \frac{1}{\lambda} \Rightarrow \lambda = \frac{1}{7{,}500}$$

$$\begin{cases} \lambda e^{-\lambda x} & 0 \le x \\ 0 & \text{سایر نقاط} \end{cases} \Rightarrow \begin{cases} \frac{1}{7{,}500} e^{-\frac{1}{7{,}500} x} & 0 \le x \\ 0 & \text{سایر نقاط} \end{cases}$$

الف)

$$P(x = 5000) = \int_{5000}^{5000} \frac{1}{7{,}500} e^{-\frac{1}{7{,}500} x}\, dx = 0$$

ب)



بنابراین فقط احتمال این محدوده را محاسبه می‌کنیم.

$$P(x \geq 10{,}000) = \int_{10{,}000}^{+\infty} \frac{1}{7{,}500}\, e^{-\frac{1}{7{,}500}x}\, dx$$

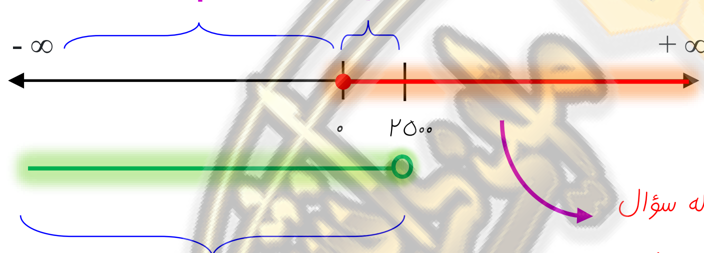
$$-e^{-\frac{1}{7{,}500}x} \Big|_{10{,}000}^{+\infty}$$

$$-\left(\underbrace{\cancel{e^{-\frac{1}{7{,}500}(+\infty)}}}_{e^{-\infty} \to 0} - e^{-\frac{1}{7{,}500}(10{,}000)} \right) = e^{-\frac{4}{3}}$$

ج)

طبق تعریف توزیع نمایی این محدوده جزءِ " سایر نقاط " است. پس احتمالش صفر است.

بنابراین فقط احتمال این محدوده را محاسبه می‌کنیم.



کل محدوده‌ای که سؤال برای آن تعریف شده

کل محدوده‌ای که سؤال خواسته

$$P(x < 2{,}500) = \underbrace{\cancel{P(x < 0)}}_{0} + P(0 \leq x < 2500)$$

$$\int_{0}^{2500} \frac{1}{7{,}500}\, e^{-\frac{1}{7{,}500}x}\, dx$$

$$-e^{-\frac{1}{7{,}500}x} \Big|_{0}^{2500}$$

$$-\left(\underbrace{\cancel{e^{-\frac{1}{7{,}500}(2500)}}}_{e^{-\frac{1}{3}}} - \underbrace{\cancel{e^{-\frac{1}{7{,}500}(0)}}}_{1} \right) = 1 - e^{-\frac{1}{3}}$$

توزیع نرمال :

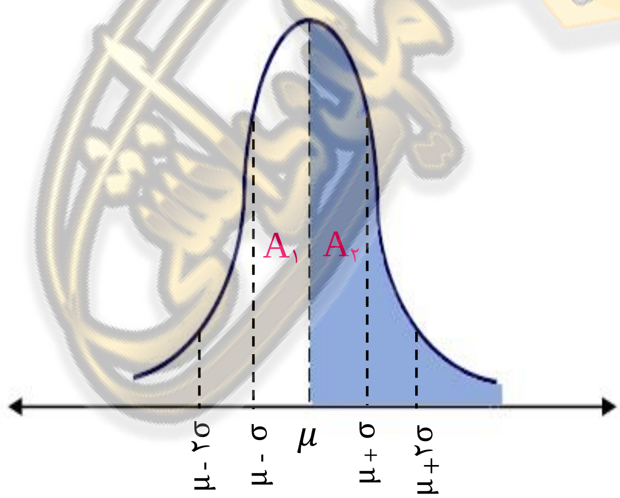
اگر x یک متغیّر تصادفی نرمال باشد، تابع چگالی احتمال آن بصورت زیر تعریف می‌شود :

$$f(x) = \frac{۱}{۲\sqrt{\pi}} e^{-\frac{(x-\mu)^{۲}}{۲\sigma^{۲}}} \qquad (-\infty < x < +\infty)$$

توزیع نرمال x را با میانگین μ و انحراف معیار σ با رابطه‌ی زیر نشان می‌دهند :

$$x \sim N = (\mu , \sigma)$$

* بدلیل تقارن توزیع نرمال بنابر شکل زیر در دو طرف میانگین، توزیع‌های سطوح $A_۱$ و $A_۲$ مساوی‌اند.



چون سطح کل زیر منحنی مساوی ۱ است پس روابط زیر برقرار است :

$A_۱ + A_۲ = ۱$

$A_۱ = A_۲ = ۰/۵$

$P(x \geq \mu) = ۰/۵$

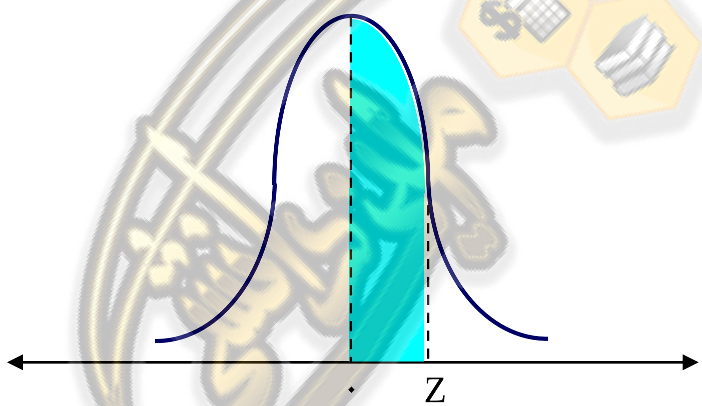
$P(x \leq \mu) = ۰/۵$

توزیع نرمال استاندارد :

اگر از متغیّر تصادفی نرمال X مقدار میانگین آن یعنی μ را کم و حاصل را به انحراف معیار X تقسیم کنیم، متغیّر جدیدی بدست می‌آید که آن را با Z نشان داده و دارای توزیع نرمال استاندارد است. این متغیّر تصادفی جدید یک متغیّر تصادفی نرمال با میانگین صفر و واریانس ۱ می‌باشد.

$$Z = \frac{X - \mu}{\sigma}$$

$$Z \sim N = (۰ , ۱)$$

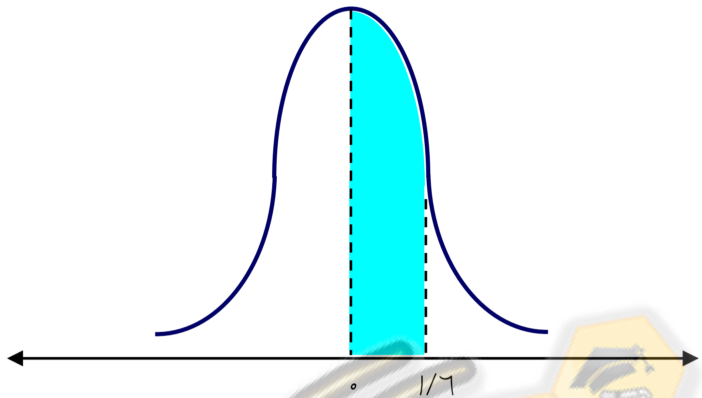


$$P(Z \geq ۰) = ۰/۵$$

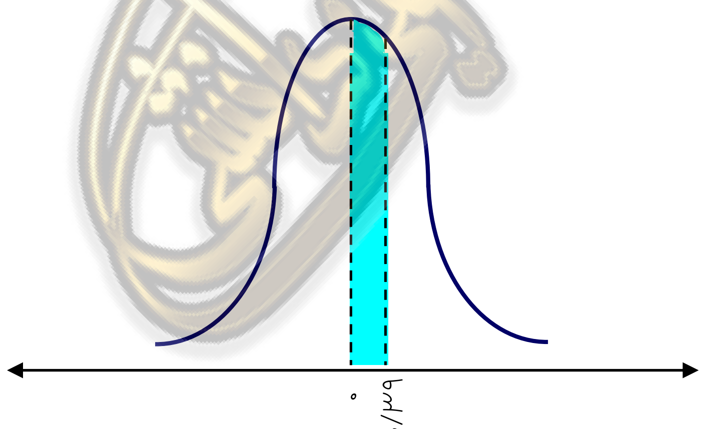
$$P(Z \leq ۰) = ۰/۵$$

مثال : با استفاده از جدول، سطح زیر منحنی نرمال هر یک از موارد زیر را محاسبه کنید :

۱) P (۰ < Z < ۱/۶) = ۰/۴۴۵۲



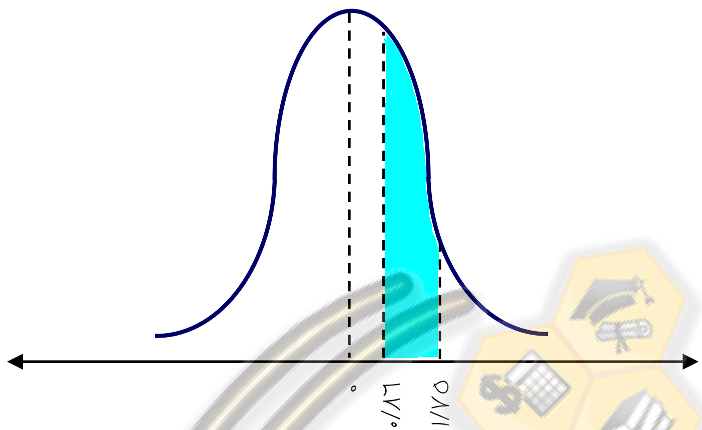
۲) P (۰ < Z < ۰/۳۹) = ۰/۱۵۱۷



۳) P (۰/۸۶ < Z < ۱/۷۵)

P (۰ < Z < ۱/۷۵) - P (۰ < Z < ۰/۸۶)

۰/۴۵۹۹ - ۰/۳۰۵۱ = ۰/۱۵۴۸

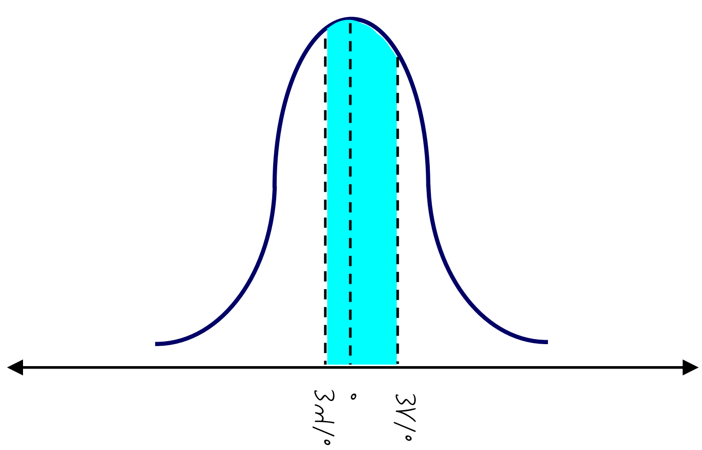


۴) P (-۰/۳۴ < Z < ۰/۸۴)

P (-۰/۳۴ < Z < ۰) + P (۰ < Z < ۰/۸۴)

P (۰ < Z < ۰/۳۴) + P (۰ < Z < ۰/۸۴)

۰/۱۳۳۱ + ۰/۲۹۹۵ = ۰/۴۳۲۶

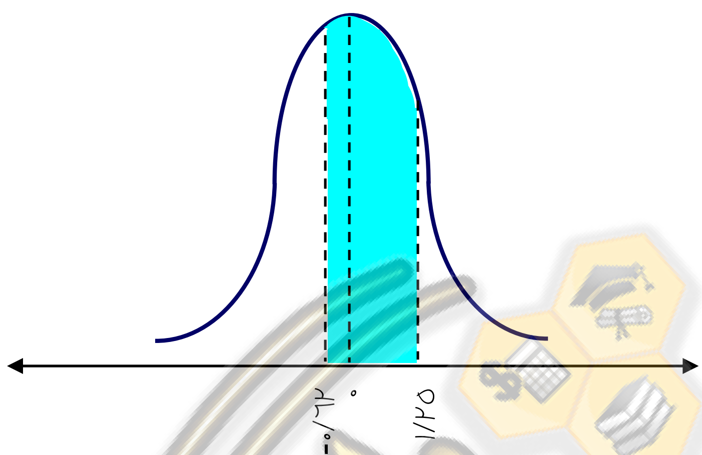


۵) $P(-۰/۶۲ < Z < ۱/۲۵)$

$P(-۰/۶۲ < Z < ۰) + P(۰ < Z < ۱/۲۵)$

$P(۰ < Z < ۰/۶۲) + P(۰ < Z < ۱/۲۵)$

$۰/۲۳۲۴ \quad + \quad ۰/۳۹۴۴ \quad = ۰/۶۲۶۸$

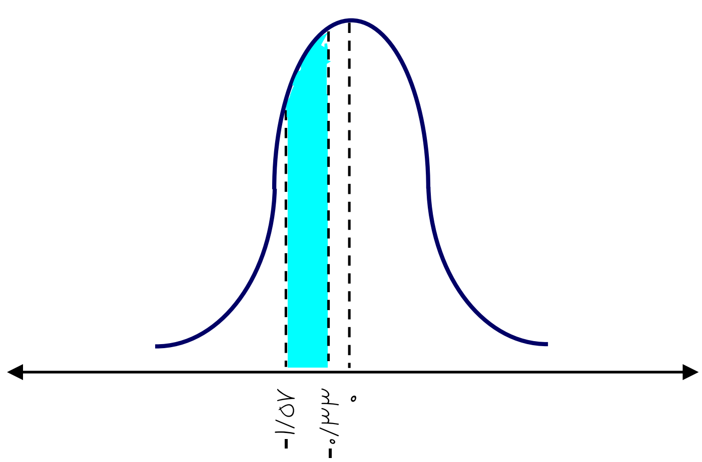


۶) $P(-۱/۵۷ < Z < -۰/۳۳)$

$P(۰/۳۳ < Z < ۱/۵۷)$

$P(۰ < Z < ۱/۵۷) - P(۰ < Z < ۰/۳۳)$

$۰/۴۴۱۸ \quad - \quad ۰/۱۲۹۳ \quad = ۰/۳۱۲۵$

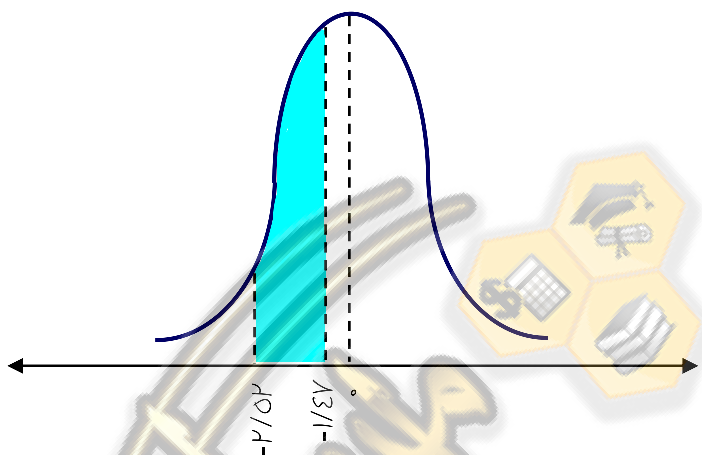


۷) P (-۲/۵۲ < Z < -۱/۴۷)

P (۱/۴۷ < Z < ۲/۵۲)

P (۰ < Z < ۲/۵۲) - P (۰ < Z < ۱/۴۷)

۰/۴۹۴۱ - ۰/۴۲۹۲ = ۰/۶۴۹



مثال : متغیّر تصادفی نرمال X دارای میانگین ۱۰ و انحراف معیار ۵ است. احتمال‌های زیر را بدست آورید :

الف) P (۱۰ < X < ۱۲)

ب) P (۱۴/۲ < X)

ج) P (۴ < X < ۱۴)

*** طبق صورت سؤال متغیّر تصادفی دارای توزیع نرمال است. از طرفی زمانی می‌توانیم از جدول توزیع احتمال نرمال استاندارد استفاده کنیم که متغیّر دارای توزیع احتمال نرمال استاندارد باشد. بنابراین با استفاده از روش زیر تک تک متغیّرها را تغییر متغیّر می‌دهیم و توزیع نرمال را به توزیع نرمال استاندارد تبدیل می‌کنیم؛ سپس از طریق جدول به جواب می‌رسیم.*

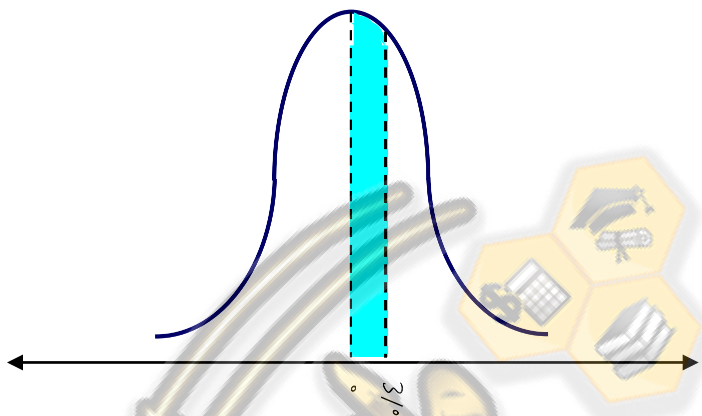
$$Z = \frac{\overset{\text{متغیّر}}{X} - \overset{\text{میانگین}}{\mu}}{\underset{\text{انحراف معیار}}{\sigma}}$$

الف)

$$P(10 < X < 12)$$

$$P\left(\frac{10-10}{5} < \frac{X-10}{5} < \frac{12-10}{5}\right)$$

$$P(0 < Z < 0/4) = 0/1554$$



ب)

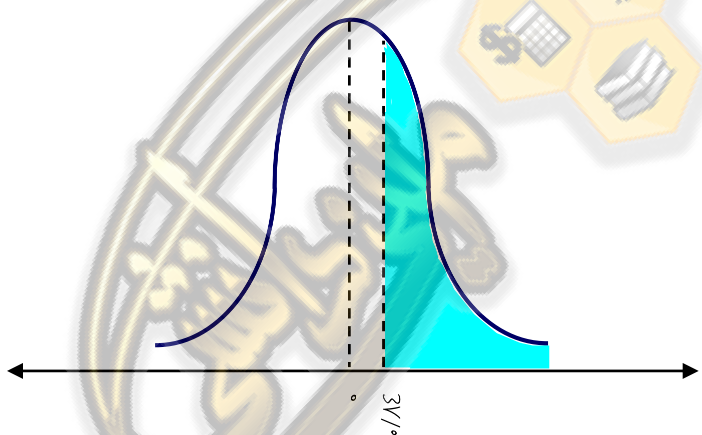
$$P\,(\,۱۴/۲ < X\,)$$

$$P\,\left(\frac{۱۴/۲ - ۱۰}{۵} < \frac{X - ۱۰}{۵}\right)$$

$$P\,(\,۰/۸۴ < Z\,)$$

$$P\,(\,۰ < Z\,) - P\,(\,۰ < Z < ۰/۸۴\,)$$

$$۰/۵ \quad - \quad ۰/۲۹۹۵ \quad = ۰/۲۰۰۵$$



ج)

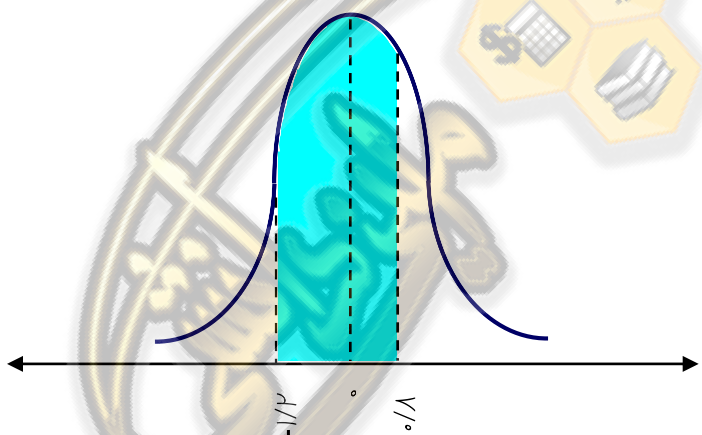
$$P\,(\,٤ < X < ١٤\,)$$

$$P\,\left(\frac{٤-١٠}{٥} < \frac{X-١٠}{٥} < \frac{١٤-١٠}{٥}\right)$$

$$P\,(\,-١/٢ < Z < ٠/٨\,)$$

$$P\,(\,٠ < Z < ٠/٨\,) + P\,(\,٠ < Z < ١/٢\,)$$

$$٠/٢٨٨١ \quad + \quad ٠/٣٨٤٩ \quad = ٠/٦٧٣٠$$



مثال : طول عمر یک نوع ماشین لباسشویی تقریباً دارای توزیع نرمال با میانگین ۳/۱ سال و انحراف معیار ۱/۲ سال است. اگر این نوع ماشین لباسشویی برای ۱ سال تضمین شده باشد، چقدر احتمال دارد که برای سال اوّل احتیاج به تعمیر پیدا کند؟

$\mu = ۳/۱$

$\sigma = ۱/۲$

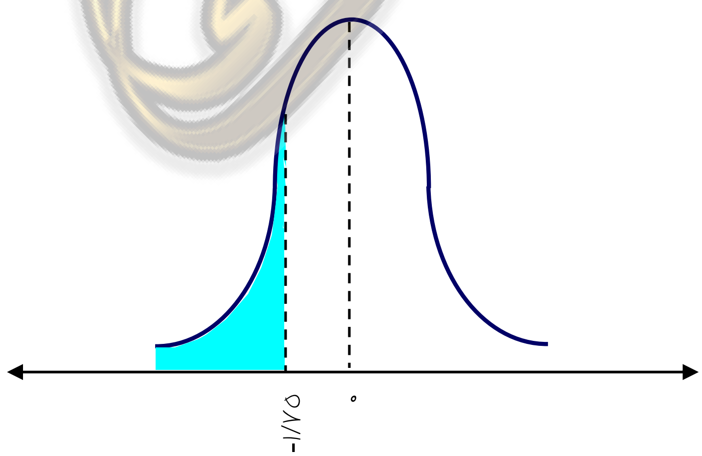
$P(X < ۱) = ?$

$$P\left(\frac{X - ۳/۱}{۱/۲} < \frac{۱ - ۳/۱}{۱/۲}\right)$$

$$P(Z < -۱/۷۵)$$

$$P(۰ < Z) - P(۰ < Z < ۱/۷۵)$$

$$۰/۵ - ۰/۴۵۹۹ = ۰/۴۰۱$$



مثال : فرض کنید که می‌خواهیم حدّاکثر تعداد افرادی را که بطور همزمان از یک آسانسور که ظرفیّت معیّنی دارد تعیین کنیم. از بررسی‌های انجام شده دریافته‌ایم که مجموع وزن ۸ نفر دارای توزیع نرمال با میانگین ۵۴۰ کیلو و انحراف معیار ۴۴/۵ کیلو است. احتمال اینکه مجموع وزن ۸ نفر از ۶۰۰ کیلو تجاوز کند چقدر است؟

$\mu = ۵۴۰$

$\sigma = ۴۴/۵$

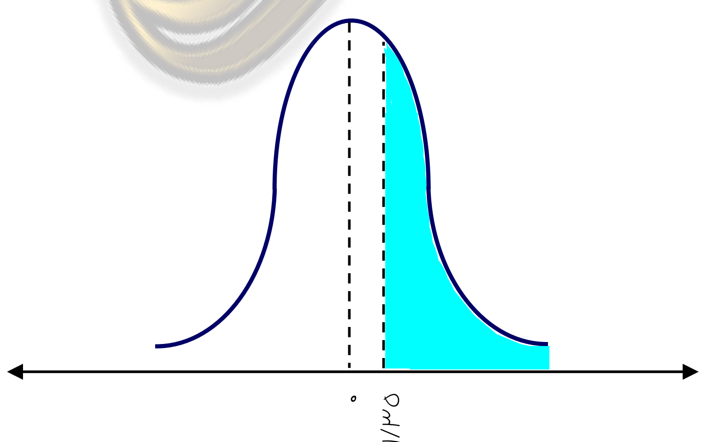
$P\ (X > ۶۰۰) =$ **?**

$$P\left(\frac{X - ۵۴۰}{۴۴/۵} > \frac{۶۰۰ - ۵۴۰}{۴۴/۵}\right)$$

$$P\ (\quad Z \quad > \quad ۱/۳۵ \quad)$$

$$P\ (۰ < Z\ ) - P\ (۰ < Z < ۱/۳۵\ )$$

$$۰/۵ \quad - \quad ۰/۴۱۱۵ \quad = ۰/۸۸۵$$



* در مواردی که n بزرگ باشد و p کوچک باشد و $np \leq ۵$ از توزیع احتمال پوآسون با $\mu = np$ برای محاسبه‌ی تقریب دو جمله‌ای استفاده می‌شود. وقتی که n بزرگ باشد ولی شرایط استفاده از توزیع پوآسون برقرار نباشد ($np > ۵$ باشد)، می‌توان از توزیع نرمال برای محاسبه‌ی تقریب احتمال دو جمله‌ای استفاده کرد.

* برای اینکه یک توزیع دو جمله‌ای تقریباً متقارن باشد، و توزیع نرمال تقریب خوبی برای آن باشد، بهتر است $n > ۲۰$ ، $np > ۵$ و $nq > ۵$ باشد. در این حالت می‌توان از توزیع نرمال با میانگین $\mu = np$ و انحراف معیار $\sigma = \sqrt{n.p.q}$ استفاده کرد.

مثال : فرض کنید که تقریباً ۲۵٪ ماشین‌های نو وانت هستند. اداره‌ی پلاک‌گذاری در یک هفته ۱۰۰ درخواست برای تعیین پلاک ماشین نو دریافت کرده است.

الف) احتمال اینکه حداّقل ۲۵ درخواست مربوط به ماشین وانت باشد، چقدر است؟

ب) احتمال اینکه حداّکثر ۱۰ درخواست مربوط به ماشین وانت باشد، چقدر است؟

$n = ۱۰۰$

$p = ۰/۲۵$

$np = ۱۰۰ \times ۰/۲۵ = ۲۵$ پوآسون نیست؛ چون ۲۵ بزرگتر از ۵ است. بنابراین از طریق نرمال حل می‌کنیم.

$$x \sim N = (\mu , \sigma) \qquad \sigma = \sqrt{n.p.q} = \sqrt{۱۰۰ \times ۰/۲۵ \times ۰/۷۵} = ۴/۳۳$$

$$x \sim N = (۲۵ , ۴/۳۳)$$

الف) $P(X \geq ۲۵)$

$$P\left(\frac{X-۲۵}{۴/۳۳} \geq \frac{۲۵-۲۵}{۴/۳۳}\right)$$

$$P(Z \geq ۰) = ۰/۵$$

ب)

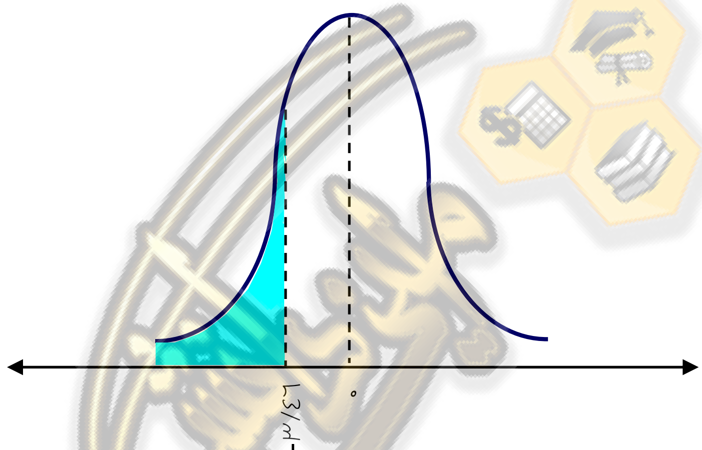
$$P(X \le ۱۰)$$

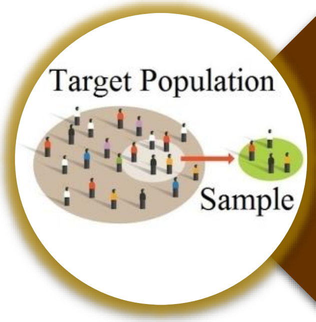
$$P\left(\frac{X - ۲۵}{٤/۳۳} \le \frac{۱۰ - ۲۵}{٤/۳۳}\right)$$

$$P(Z \le -۳/٤٦)$$

$$P(۰ < Z) - P(۰ < Z < ۳/٤٦)$$

$$۰/۵ \quad - \quad ۰/٤۹۹۷ \quad = ۰/۰۰۰۳$$





فصل سوّم

نمونه‌گیری و آماره‌ها

آماره :

متغیّر تصادفی که مقادیر آن با استفاده از داده‌های موجود در نمونه محاسبه می‌شود، آماره نام دارد. میانگین نمونه‌ای و واریانس نمونه‌ای دو آماره‌ی مهم هستند. توزیع نمونه‌ای توزیع احتمال آماره‌ای که با استفاده از داده‌های یک نمونه‌ی تصادفی با اندازه‌ی n از جامعه محاسبه می‌شود، توزیع نمونه‌ای آن آماره نامیده می‌شود. در تعریف توزیع میانگین نمونه‌ای دو نوع جامعه‌ی متناهی و نامتناهی را در نظر می‌گیریم.

جامعه‌ی متناهی :

اگر جامعه‌ای متناهی با N عضو دارای میانگین μ و واریانس σ^2 داشته باشیم، و نمونه‌ای با اندازه‌ی n بطور تصادفی از این جامعه استخراج کنیم، در اینصورت میانگین ($\mu_{\bar{x}}$) و خطای معیار ($d_{\bar{x}}$) بشرح زیر خواهد بود :

$$\mu_{\bar{x}} = E(\bar{x}) = \mu$$

$$d_{\bar{x}} = \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}}$$

جامعه‌ی نامتناهی :

وقتی جامعه نامتناهی باشد یا اندازه‌ی جامعه نسبت به اندازه‌ی نمونه بزرگ باشد، مقدار $\sqrt{\frac{N-n}{N-1}}$ تقریباً برابر با ۱ است؛ پس میانگین و خطای معیار از رابطه‌ی زیر بدست می‌آید :

$$\mu_{\bar{x}} = E(\bar{x}) = \mu$$

$$d_{\bar{x}} = \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

* مقدار $\sqrt{\frac{N-n}{N-1}}$ را ضریب تصحیح جامعه‌ی متناهی گویند. وقتی مقدار این ضریب از ۰/۹۵ بیشتر باشد، یا بطور معادل $n \leq 0/05\ N$ جامعه را نامتناهی فرض کرده و این ضریب را در محاسبه‌ی انحراف معیار ۱ فرض می‌کنیم.

قضیه حد مرکزی :

اگر نمونه‌ی تصادفی به اندازه‌ی n از جامعه‌ای غیر نرمال با میانگین μ و انحراف معیار σ استخراج شده باشد و n بزرگ باشد، توزیع نمونه‌ای $\bar{x}$ تقریباً نرمال است و هرچه n بزرگتر باشد، تقریب بهتر خواهد بود. میانگین $\bar{x}$ برابر با میانگین جامعه آماری (μ) و انحراف معیار $\bar{x}$ با توجّه به اندازه‌ی جامعه از روابط فوق بدست می‌آید.

* شرایط تقریب خوب برای مقدار n تا توزیع نمونه‌ای $\bar{x}$ تقریباً نرمال باشد، به اینصورت است :

۱. اگر توزیع فراوانی داده‌های نمونه تقریباً متقارن باشد، اگر $n \geq 30$
۲. اگر توزیع فراوانی داده‌های نمونه چوله باشد. اگر $n \geq 100$

مثال : جامعه‌ای دارای ۶ عنصر با مقادیر ۶ , ۵ , ۱۰ , ۱۳ , ۱۴ , ۱۸ است. اگر نمونه‌ی تصادفی با اندازه‌ی ۳ از این جامعه استخراج کنیم، میانگین و انحراف معیار $\bar{x}$ را بدست آورید.

$$\mu = \frac{\sum_{i=۱}^{N} x_i}{N}$$

$$\mu = \frac{\sum_{i=۱}^{۶} x_i}{۶} = \frac{۶ + ۵ + ۱۰ + ۱۳ + ۱۴ + ۱۸}{۶} = \frac{۶۶}{۶} = ۱۱$$

$$\sigma_x^{۲} = \frac{\sum_{i=۱}^{N} (x_i - \mu)^{۲}}{N} = \frac{(۵-۱۱)^{۲} + (۶-۱۱)^{۲} + (۱۰-۱۱)^{۲} + (۱۳-۱۱)^{۲} + (۱۴-۱۱)^{۲} + (۱۸-۱۱)^{۲}}{۶}$$

$$= \frac{۳۶ + ۲۵ + ۱ + ۴ + ۹ + ۴۹}{۶} = ۲۰/۶$$

$$\sigma_x = \sqrt{۲۰/۶} = ۴/۵$$

$$\begin{cases} \mu_{\bar{x}} = \mu = ۱۱ \\[2ex] d_{\bar{x}} = \sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}} \times \sqrt{\dfrac{N-n}{N-۱}} \\[2ex] \dfrac{۴/۵}{\sqrt{۳}} \times \sqrt{\dfrac{۶-۳}{۶-۱}} \sim ۲ \end{cases}$$

*اگر جامعه‌ای دارای توزیع نرمال با میانگین μ و انحراف معیار σ باشد و نمونه‌ی تصادفی به اندازه‌ی n از آن استخراج شده باشد، در اینصورت میانگین نمونه دارای توزیع نرمال با میانگین μ و انحراف معیار $\frac{\sigma}{\sqrt{n}}$ است.

مثال : X طول میله‌های فولادی تولید شده توسّط یک کارخانه، دارای میانگین " $\mu = ۶$ متر " و واریانس " $\sigma^۲ = ۶۴$ سانتی متر " است. این میله‌ها در دسته‌های ۴۰تایی بسته‌بندی می‌شوند. اگر توزیع طول میله‌ها تقریباً متقارن باشد،

الف) احتمال اینکه میانگین طول میله‌های یک بسته که بطور تصادفی انتخاب شده کمتر از ۵۹۸ سانتی متر باشد، چقدر است ؟

ب) احتمال اینکه میانگین طول میله‌های یک بسته که بطور تصادفی انتخاب شده کمتر از ۵۹۸ سانتی متر یا بیشتر از ۶۰۱ سانتی متر باشد، چقدر است ؟

$n = ٤٠ > ٣٠$ ، متقارن ، تقریباً نرمال

$\mu = ٦٠٠ \longrightarrow \mu_{\bar{x}} = ٦٠٠$

$\sigma^٢ = ٦٤ \longrightarrow \sigma = ٨ \longrightarrow \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{٨}{\sqrt{٤٠}} = ١/٢$

$$\bar{x} \sim N = \left(\mu , \frac{\sigma}{\sqrt{n}}\right)$$

$$\bar{x} \sim N = (٦٠٠ , ١/٢)$$

الف)

$$P(\bar{x} < ٥٩٨)$$

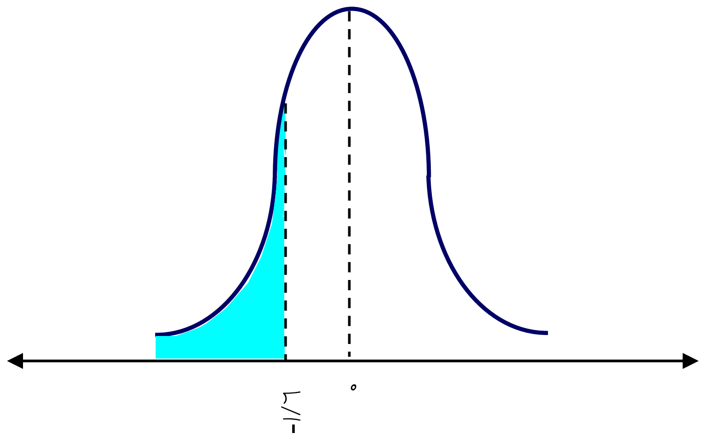
$$P\left(\frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} < \frac{٥٩٨ - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right)$$

$$P\left(\frac{\bar{x} - ٦٠٠}{١/٢} < \frac{٥٩٨ - ٦٠٠}{١/٢}\right)$$

$$P(\quad Z \quad < \quad -١/٦ \quad)$$

$$P(٠ < Z) - P(٠ < Z < ١/٦)$$

$$٠/٥ \quad - \quad ٠/٤٤٥٢ \quad = ٠/٠٥٤٨$$



ب)

$$P(۵۹۸ < \bar{x} < ۶۰۰)$$

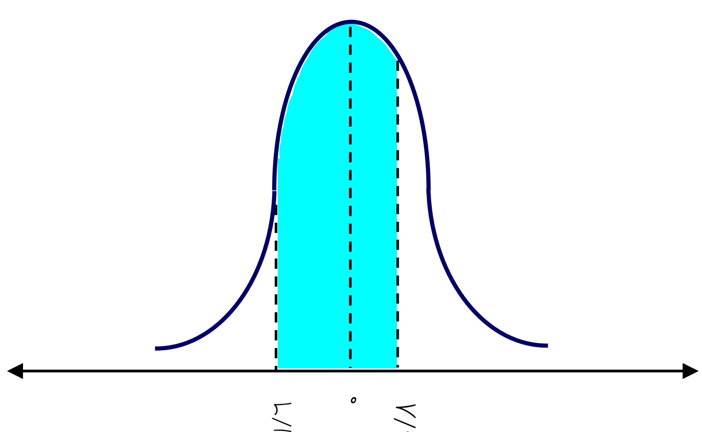
$$P\left(\frac{۶۰۱ - \mu_{\bar{x}}}{\sigma_{\bar{x}}} < \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} < \frac{۵۹۸ - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right)$$

$$P\left(\frac{۶۰۱ - ۶۰۰}{۱/۲} < \frac{\bar{x} - ۶۰۰}{۱/۲} < \frac{۵۹۸ - ۶۰۰}{۱/۲}\right)$$

$$P(۰/۸ < \bar{Z} < -۱/۶)$$

$$P(۰ < \bar{Z} < ۰/۸) + P(۰ < \bar{Z} < ۱/۶)$$

$$۰/۲۸۸۱ + ۰/۴۴۵۲ = ۰/۷۳۳۳$$



مثال : از مصرف‌کنندگان نوعی محصول خواسته شده که کیفیّت محصول شرکت را با دادن ۱ تا ۴ امتیاز رتبه‌بندی کنند. میانگین و انحراف جامعه به ترتیب عبارت است از ۲/۳ و ۰/۷۵ می‌باشد. برای نمونه تصادفی مرکّب از ۱۵۰ مصرف‌کننده.

الف) احتمال اینکه میانگین نمونه از ۲/۱۵ تجاوز کند، چقدر است؟

ب) احتمال اینکه در فاصله‌ی ۰/۱۵± از μ قرار بگیرد، چقدر است؟

$n = ۱۵۰ > ۳۰$ ، متقارن ، تقریباً نرمال

$$\mu = ۲/۳ \longrightarrow \mu_{\bar{x}} = ۲/۳$$

$$\sigma = ۰/۷۵ \longrightarrow \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{۰/۷۵}{\sqrt{۱۵۰}} = ۰/۰۶$$

$$\bar{x} \sim N = \left(\mu , \frac{\sigma}{\sqrt{n}}\right)$$

$$\bar{x} \sim N = (۲/۳ , ۰/۰۶)$$

الف)

$$P(\bar{x} > ۲/۱۵)$$

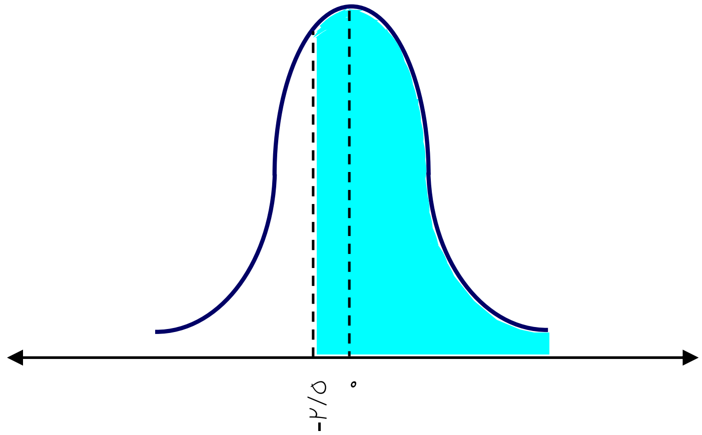
$$P\left(\frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} > \frac{۲/۱۵ - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right)$$

$$P\left(\frac{\bar{x} - ۲/۳}{۰/۰۶} > \frac{۲/۱۵ - ۲/۳}{۰/۰۶}\right)$$

$$P(\bar{Z} > -۲/۵)$$

$$P(۰ < \bar{Z}) + P(۰ < \bar{Z} < ۲/۵)$$

$$۰/۵ + ۰/۴۹۳۸ = ۰/۹۹۳۸$$



ب)

$$P(2/15 < \bar{x} < 2/45)$$

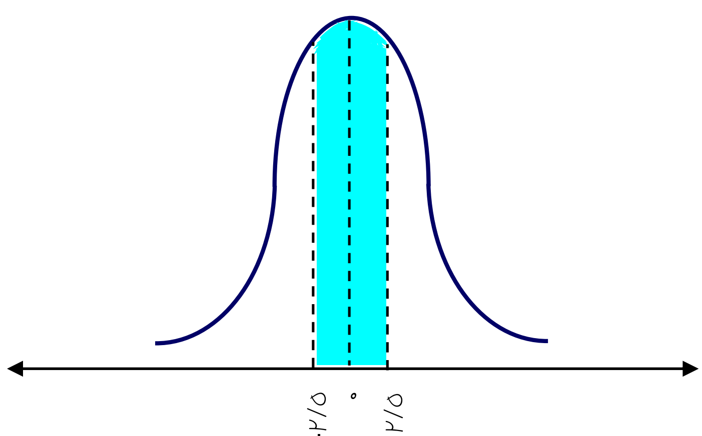
$$P\left(\frac{2/15-\mu_{\bar{x}}}{\sigma_{\bar{x}}} < \frac{\bar{x}-\mu_{\bar{x}}}{\sigma_{\bar{x}}} < \frac{2/45-\mu_{\bar{x}}}{\sigma_{\bar{x}}}\right)$$

$$P\left(\frac{2/15-2/3}{0/06} < \frac{\bar{x}-2/3}{0/06} < \frac{2/45-2/3}{0/06}\right)$$

$$P(-2/5 < \bar{Z} < 2/5)$$

$$P(0 < \bar{Z} < 2/5) \times 2$$

$$0/4938 \times 2 = 0/9876$$



مثال : از یک جامعه‌ی بزرگ فروشگاه‌های لوازم خانگی برای برآورد میانگین ضرر ناشی از سرقت در هر فروشگاه در طول ماه نمونه‌گیری می‌شود. میانگین و انحراف معیار جامعه برحسب تومان عبارت است از : $\mu = ۳۸۰۰$ و $\sigma = ۲۱۰۰$. برای نمونه تصادفی به اندازه‌ی ۲۵۰ فروشگاه احتمال اینکه میانگین نمونه در فاصله ±۱۵۰ از میانگین جامعه قرار گیرد، چقدر است؟

$$n = ۲۵۰ > ۳۰$$

$$\mu = ۳۸۰۰ \longrightarrow \mu_{\bar{x}} = ۳۸۰۰$$

$$\sigma = ۲۱۰۰ \longrightarrow \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{۲۱۰۰}{\sqrt{۲۵۰}} = ۱۳۲/۸$$

$$\bar{x} \sim N = \left(\mu , \frac{\sigma}{\sqrt{n}}\right)$$

$$\bar{x} \sim N = (۳۸۰۰ , ۱۳۲/۸)$$

$$P(۳۶۵۰ < \bar{x} < ۳۹۵۰)$$

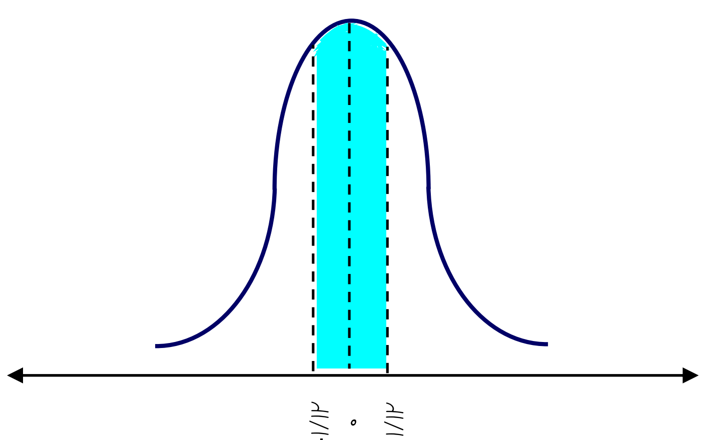
$$P\left(\frac{۳۶۵۰ - \mu_{\bar{x}}}{\sigma_{\bar{x}}} < \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} < \frac{۳۹۵۰ - \mu_{\bar{x}}}{\sigma_{\bar{x}}}\right)$$

$$P\left(\frac{۳۶۵۰ - ۳۸۰۰}{۱۳۲/۸} < \frac{\bar{x} - ۳۸۰۰}{۱۳۲/۸} < \frac{۳۹۵۰ - ۳۸۰۰}{۱۳۲/۸}\right)$$

$$P(-۱/۱۲ < \bar{Z} < ۱/۱۲)$$

$$P(۰ < \bar{Z} < ۱/۱۲) \times ۲$$

$$۰/۳۶۸۶ \times ۲ = ۰/۷۳۷۲$$



توزیع نسبت نمونه‌ای :

اگر نسبت افراد یک جامعه را که دارای صفت مشخّصی هستند با p نشان دهیم و مقدار p معلوم نباشد، در اینصورت برای نتیجه‌گیری درباره‌ی مقدار واقعی p ابتدا نمونه‌ای به اندازه‌ی n از جامعه انتخاب و نسبت افرادی را که دارای مشخّصه‌ی مورد نظرند در این نمونه بدست می‌آوریم. نسبت بدست آمده را نسبت نمونه‌ای گویند و با $\bar{p} = \dfrac{X}{n}$ نمایش می‌دهند که X تعداد افراد دارای مشخّصه‌ی مورد نظر یا تعداد موفّقیّت‌های مشاهده شده در نمونه است.

* توزیع X دارای توزیع دوجمله‌ای با پارامتر p و n است. اگر n بزرگ باشد، X دارای توزیع نرمال با میانگین $\mu = n.p$ و انحراف معیار $\sigma = \sqrt{n.p.q}$ است.

* توزیع $\bar{p}$ اگر $\bar{p} = \dfrac{X}{n}$ را بعنوان تبدیل متغیّری از X درنظر بگیریم و n بزرگ باشد، $\bar{p}$ دارای توزیع تقریبی نرمال با میانگین $\mu_{\bar{p}} = p$ و انحراف معیار $\sigma_{\bar{p}} = \sqrt{\dfrac{p.q}{n}}$ است.

مثال : اگر نمونه‌ای به اندازه‌ی $n = ۲۰۰$ بطور تصادفی از جامعه‌ای انتخاب شود که نسبت موفّقیّت‌ها در جامعه $p = ۰/۷$ باشد؛ احتمال اینکه اختلاف نسبت نمونه‌ای و نسبت جامعه کمتر از ۰/۰۵ باشد، چقدر است؟

$n = 200$

$p = 0/7 \longrightarrow \mu_{\bar{p}} = 0/7$

$$\sigma_{\bar{p}} = \sqrt{\frac{p.q}{n}} = \sqrt{\frac{0/7 \times 0/3}{200}} = 0/032$$

$$P(\,|\bar{p} - p| < 0/05\,)$$

$$P(-0/05 < \bar{p} - p < 0/05)$$

$$P(-0/05 < \bar{p} - 0/7 < 0/05)$$

$$P(-0/05 + 0/7 < \bar{p} < 0/05 + 0/7)$$

$$P(0/65 < \bar{p} < 0/75)$$

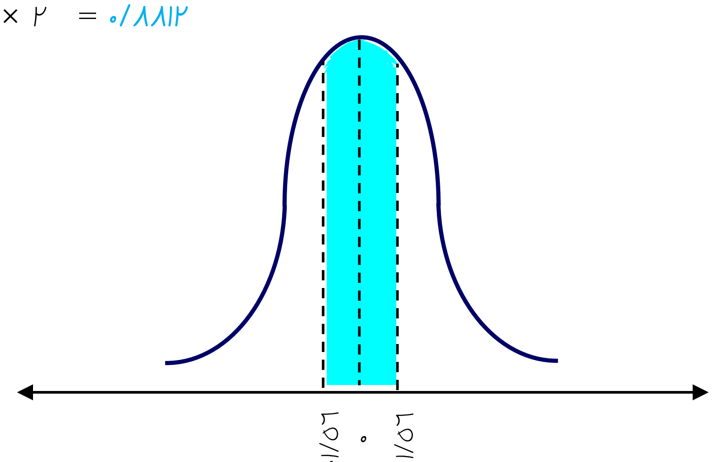
$$P\left(\frac{0/65 - \mu_{\bar{p}}}{\sigma_{\bar{p}}} < \frac{\bar{p} - \mu_{\bar{p}}}{\sigma_{\bar{p}}} < \frac{0/75 - \mu_{\bar{p}}}{\sigma_{\bar{p}}}\right)$$

$$P\left(\frac{0/65 - 0/7}{0/032} < \frac{\bar{p} - 0/7}{0/032} < \frac{0/75 - 0/7}{0/032}\right)$$

$$P(\quad -1/56 \quad < \quad \bar{Z} \quad < \quad 1/56 \quad)$$

$$P(0 < \bar{Z} < 1/56) \times 2$$

$$0/4406 \quad \times 2 \quad = 0/8812$$



مثال : بررسی‌های گذشته نشان داده است که ۶۳٪ از زنانی که فرزندان زیر ۱۸ سال دارند در خارج از منزل کار می‌کنند. برای تعیین صحّت این ادّعا نمونه‌ی تصادفی از ۱۰۰ زن که دارای فرزندان زیر ۱۸ سال هستند انتخاب و مشاهده شده که حدّاقل ۸۴ نفر در خارج از منزل کار می‌کنند.

الف) اگر نسبت گزارش شده درست باشد، درباره‌ی توزیع $\bar{p}$ توضیح دهید.

ب) احتمال اینکه در یک نمونه‌ی تصادفی مقدار $\bar{p} \geq ۰/۸۴$ باشد را بیابید.

الف)

$n = ۱۰۰$

$p = ۰/٦٣ \longrightarrow \mu_{\bar{p}} = ۰/٦٣$

$q = ۰/٣٧$

$$\sigma_{\bar{p}} = \sqrt{\frac{p.q}{n}} = \sqrt{\frac{۰/٦٣ \times ۰/٣٧}{۱۰۰}} = ۰/۰۴۷$$

$\bar{p} \sim N = (۰/٦٣ \ , \ ۰/۰٤٧)$

ب)

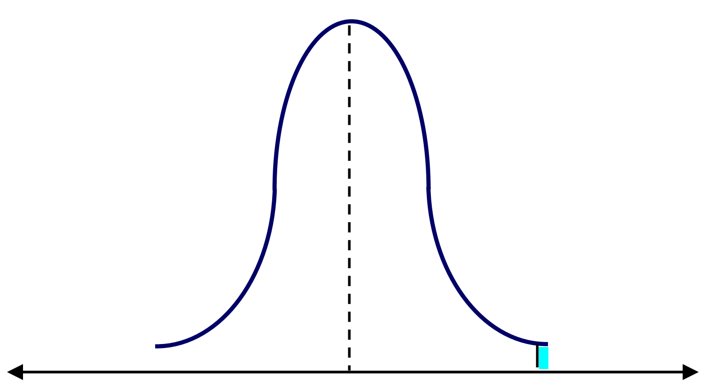
$P(۰ < ۸٤)$

$P(۰/۸٤ \leq \bar{p})$

$$P\left(\frac{۰/٦٥ - \mu_{\bar{p}}}{\sigma_{\bar{p}}} < \frac{\bar{p} - \mu_{\bar{p}}}{\sigma_{\bar{p}}}\right)$$

$$P\left(\frac{۰/۸٤ - ۰/٦٣}{۰/۰٤٧} < \frac{\bar{p} - ۰/٦٣}{۰/۰٤٧}\right)$$

$$P(٤/٤٦ < \bar{Z}) \cong ۰$$



۰ ٤/٤٦

فصل چهارم

برآورد میانگین جامعه

پارامتر : عدد یا متغیّری از جامعه است که خصوصیّتی از آن را توصیف می‌کند؛ مانند میانگین، واریانس، میانه و ... .

آماره : اصطلاحی است که درمورد نمونه استفاده می‌شود و خصوصیّتی از آن را بررسی می‌کند.

برآورد : برای بدست آوردن پارامتر مجهول جامعه با توجّه به اطّلاعات بدست آمده‌ی حاصل از نمونه به ۲ صورت عمل می‌شود :

۱. برآورد یا تخمین
۲. استفاده از آزمون فرض پارامتر مجهول

برآورد یک پارامتر مجهول به ۲ صورت انجام می‌شود.

A. **برآورد نقطه‌ای :** وقتی مشخّصه‌ای از جامعه با یک عدد برآورد شود، آن عدد را برآورد نقطه‌ای آن مشخّصه می‌نامیم.

برآوردکننده : آماره‌ای است که برای برآورد کردن مشخّصه‌ای از جامعه بکار می‌رود. اگر پارامتر جامعه θ باشد، برآوردکننده‌ی پارامتر θ را با $\hat{\theta}$ نمایش می‌دهند.

ویژگی‌های برآوردکننده‌های نقطه‌ای خوب :

I. نا اُریب (مستقیم) : برآوردکننده‌ای را نا اُریب گویند که میانگین توزیع نمونه‌ای آن با پارامتر مورد برآورد جامعه برابر باشد. بعبارتی اگر $\hat{\theta}$ یک برآوردکننده‌ی نقطه‌ای از پارامتری مجهول (θ) باشد، این برآوردکننده را نا اُریب گویند؛ اگر $E(\hat{\theta}) = \theta$ باشد.

II. کارآیی : اگر دو برآوردکننده‌ی نا اُریب برای پارامتر θ موجود باشد، برآوردکننده‌ای که واریانس کوچکتری داشته باشد، کارآیی نسبی بیشتری نسبت به دیگری دارد.

III. سازگاری : برآوردکننده‌ای مانند $\hat{\theta}$ را یک برآوردکننده‌ی سازگار برای پارامتر θ می‌نامیم، هرگاه با افزایش n، $\hat{\theta}$ با احتمال بیشتری به θ نزدیک شود.

B. **برآورد فاصله‌ای :** در این روش با استفاده از برآورد کننده‌ی نقطه‌ایِ پارامتر جامعه، حدودی برای پارامتر مورد نظر پیدا می‌کنیم. این حدود یک بازه یا یک فاصله را مشخّص می‌کند که ممکن است مقدار واقعی پارامتر را دربرداشته باشد. این بازه را فاصله‌ی اطمینان یا برآورد فاصله‌ای پارامتر گویند و بصورت زیر نمایش می‌دهیم :

$$L < \theta < U$$

کران پایین

کران بالا

پارامتر مجهول

ضریب اطمینان : با C نمایش داده می‌شود و مقدارش برابر با احتمال اینکه فاصله‌ی اطمینان شامل مقدار واقعی پارامتر برآورد شده باشد. ضریب اطمینان را بصورت درصد بیان می‌کنند.

$$C = P\,(L < \theta < U)$$

خطای برآورد : درحالت کلّی، برآورد، مقدار دقیق پارامتر را معلوم نمی‌کند و دارای مقداری خطاست اگر خطای برآورد را با d نمایش دهیم، مقدار آن از رابطه‌ی زیر بدست می‌آید :

d = | مقدار واقعی پارامتر - مقدار برآورد پارامتر |

خطای برآورد μ برای نمونه‌های بزرگ بصورت زیر است :

$$d = \left| \bar{x} - \mu \right|$$

برآورد نقطه‌ای میانگین جامعه :

قضیه : بهترین برآوردکننده میانگین جامعه ($\bar{x}$) میانگین نمونه‌ای است که درصورت بزرگ بودن اندازه‌ی نمونه‌ی جامعه، دارای توزیع تقریبی نرمال با میانگین μ و انحراف معیار $\frac{\sigma}{\sqrt{n}}$ است. بنابراین چون μ نامعلوم است، خطای برآورد نیز نامعلوم است ولی اگر اندازه‌ی نمونه بزرگ باشد می‌توانیم یک عبارت احتمالی برای خطای برآورد بصورت زیر بنویسیم :

$$P\left(d < z_{\frac{\alpha}{۲}} \cdot \sigma_{\bar{x}}\right) = ۱ - \alpha$$

که α عددی بین صفر و ۱ و $z_{\frac{\alpha}{۲}}$ نقطه‌ی ٪$\frac{\alpha}{۲}$ توزیع نرمال است. بنابراین رابطه‌ی فوق با ٪$(۱ - \alpha)۱۰۰$ اطمینان، حکم می‌کنیم که خطای برآورد کمتر از $(z_{\frac{\alpha}{۲}} \cdot \sigma_{\bar{x}})$ است.

برآورد نقطه‌ای واریانس جامعه بهترین برآورد کننده‌ی $\sigma^۲$ است. اگر $\sigma^۲$ معلوم نباشد، $S^۲$ واریانس نمونه‌ای آماره‌ای است که ویژگی‌های یک برآوردکننده‌ی خوب را دارد.

برآورد فاصله‌ی μ برای نمونه‌های بزرگ :

فاصله‌ی اطمینان μ برای نمونه‌های بزرگ درصورت معلوم بودن یا نبودن انحراف معیار جامعه بصورت زیر بدست می‌آید :

۱. اگر انحراف معیار جامعه معلوم باشد :

$$d = z_{\frac{\alpha}{۲}} \cdot \frac{\sigma}{\sqrt{n}}$$

$$L = \bar{x} - z_{\frac{\alpha}{۲}} \cdot \frac{\sigma}{\sqrt{n}}$$

$$V = \bar{x} + z_{\frac{\alpha}{۲}} \cdot \frac{\sigma}{\sqrt{n}}$$

۲. اگر انحراف معیار جامعه معلوم نباشد :

$$d = z_{\frac{\alpha}{۲}} \cdot \frac{S}{\sqrt{n}}$$

$$L = \bar{x} - z_{\frac{\alpha}{۲}} \cdot \frac{S}{\sqrt{n}}$$

$$V = \bar{x} + z_{\frac{\alpha}{۲}} \cdot \frac{S}{\sqrt{n}}$$

برآورد فاصله‌ی μ برای نمونه‌های کوچک :

وقتی اندازه‌ی نمونه کوچک باشد، توزیع نمونه‌ای $\bar{x}$ دیگر نرمال نخواهد بود و به توزیع جامعه بستگی دارد. در این حالت حدود اطمینان را برای جامعه‌ی نرمال بدست می‌آوریم. وقتی جامعه نرمال باشد، برای نمونه‌ی تصادفی به اندازه‌ی n با میانگین $\bar{x}$ و انحراف معیار S، متغیّر تصادفی $\dfrac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$ دارای توزیع t با $(n - ۱)$ درجه آزادی است.

در اینصورت حدود اطمینان μ با ضریب اطمینان $(۱ - \alpha)$ وقتی $n < ۳۰$ باشد، عبارت است از :

$$L = \bar{x} - t_{(\frac{\alpha}{۲}, n-۱)} \cdot \frac{S}{\sqrt{n}}$$

$$V = \bar{x} + t_{(\frac{\alpha}{۲}, n-۱)} \cdot \frac{S}{\sqrt{n}}$$

که $t_{(\frac{\alpha}{۲}, n-۱)}$ نقطه‌ی ٪ $\frac{\alpha}{۲}$ توزیع t با $(n - ۱)$ درجه آزادی است. رابطه‌ی فوق برای زمانی است که σ نامعلوم باشد؛ امّا اگر σ معلوم باشد، متغیّر تصادفی $\dfrac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$ دارای توزیع نرمال استاندارد خواهد بود و حدود اطمینان μ از رابطه‌ی زیر بدست می‌آید :

$$L = \bar{x} - z_{\frac{\alpha}{۲}} \cdot \frac{\sigma}{\sqrt{n}}$$

$$V = \bar{x} + z_{\frac{\alpha}{۲}} \cdot \frac{\sigma}{\sqrt{n}}$$

حدود اطمینان μ در جامعه‌ی غیرنرمال :

حدود اطمینان را وقتی جامعه دقیقاً نرمال نباشد، با ضریب اطمینانِ دقیقاً $۱ - \alpha$ نمی‌توان بیان کرد، بلکه تقریباً $۱ - \alpha$ بیان می‌کنند و از رابطه‌ی زیر بدست می‌آید :

$$L = \bar{x} - t_{(\frac{\alpha}{۲}, n-۱)} \cdot \frac{\sigma}{\sqrt{n}}$$

$$V = \bar{x} + t_{(\frac{\alpha}{۲}, n-۱)} \cdot \frac{\sigma}{\sqrt{n}}$$

*

$t_{/۰۵, ۱۳} = ۱/۷۷۱$	$z_{/۰۵} = ۱/۶۴۵$
$t_{/۰۲۵, ۱۵} = ۲/۱۳۱$	$z_{/۰۰۵} = ۲/۵۷۶$
	$z_{/۰۲۵} = ۱/۹۶$

مثال : نمونه‌ی تصادفی به اندازه‌ی $n = ۱۰۰$ از جامعه‌ی نامتناهی با میانگین μ و واریانس $\sigma^۲$ استخراج شده است. اگر بخواهیم با ۹۵٪ اطمینان قضاوت کنیم، حدّاکثر خطای برآورد μ برای مقادیر مختلف σ که در زیر داده شده، چقدر است؟

الف) $\sigma = ۱۰$

ب) $\sigma = ۱۰۰$

$$1 - \alpha = 95\% \longrightarrow \alpha = 0/05 \longrightarrow \frac{\alpha}{2} = 0/025$$

الف)

$$d = z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

$$d = z_{0/025} \cdot \frac{10}{\sqrt{100}}$$

$$d = 1/96 \times \frac{10}{\sqrt{100}} = 1/96$$

ب)

$$d = z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

$$d = z_{0/025} \cdot \frac{100}{\sqrt{100}}$$

$$d = 1/96 \times \frac{100}{\sqrt{100}} = 19/6$$

مثال : یک نمونه‌ی تصادفی ۱۰۰تایی از کارمندان یک شرکت انتخاب و حقوق ماهیانه آنها را ثبت کردیم. میانگین و انحراف معیار حقوق‌ها بشرح زیر است.

$\bar{x} = ۱$

$S = ۲۳۹$

یک فاصله اطمینان ۹۵ درصدی برای میانگین حقوق کارمندان بسازید :

$$۱ - \alpha = ۹۵٪ \longrightarrow \alpha = ۰/۰۵ \longrightarrow \frac{\alpha}{۲} = ۰/۰۲۵$$

$$L < \mu < V$$

$$\bar{x} - z_{\frac{\alpha}{۲}} \cdot \frac{S}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{۲}} \cdot \frac{S}{\sqrt{n}}$$

$$۱ - z_{۰/۰۲۵} \cdot \frac{S}{\sqrt{n}} < \mu < ۱ + z_{۰/۰۲۵} \cdot \frac{S}{\sqrt{n}}$$

$$۱ - ۱/۹۶ \cdot \frac{۲۳۹}{\sqrt{۱۰۰}} < \mu < ۱ + ۱/۹۶ \cdot \frac{۲۳۹}{\sqrt{۱۰۰}}$$

$$-۴۵/۸۴ < \mu < ۴۷/۸۴$$

مثال : مبالغ هزینه پستی برای نمونه‌ای به اندازه‌ی $n = ۴۰۰$ بسته در روزی خاص که بوسیله‌ی اداره‌ی پست جابجا شده‌اند، ثبت شده است. میانگین و انحراف معیار نمونه‌ای عبارت است از :

$\bar{x} = ۳۶۴/۷$

$S = ۲۵۹$

الف) انحراف معیار $\bar{x}$ را برآورد کنید.

ب) یک فاصله اطمینان ۹۵ درصدی برای میانگین بیابید.

ج) یک فاصله اطمینان ۹۰ درصدی برای میانگین بیابید.

الف)

$$\sigma_{\bar{x}} = \frac{S}{\sqrt{n}} = \frac{۲۵۹}{\sqrt{۴۰۰}} = ۱۲/۹۵$$

ب)

$$۱ - \alpha = ۹۵\% \longrightarrow \alpha = ۰/۰۵ \longrightarrow \frac{\alpha}{۲} = ۰/۰۲۵$$

$$L < \mu < V$$

$$\bar{x} - z_{\frac{\alpha}{۲}} \cdot \frac{S}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{۲}} \cdot \frac{S}{\sqrt{n}}$$

$$۳۶۴/۷ - z_{۰/۰۲۵} \cdot \frac{S}{\sqrt{n}} < \mu < ۳۶۴/۷ + z_{۰/۰۲۵} \cdot \frac{S}{\sqrt{n}}$$

$$۳۶۴/۷ - ۱/۹۶ \times (۱۲/۹۵) < \mu < ۳۶۴/۷ + ۱/۹۶ \times (۱۲/۹۵)$$

$$۳۳۹/۳۱۸ < \mu < ۳۹۰/۰۸۲$$

ج)

$$\text{۱} - \alpha = \text{۹۰٪} \longrightarrow \alpha = \text{۰/۱} \longrightarrow \frac{\alpha}{\text{۲}} = \text{۰/۰۵}$$

$$L < \mu < V$$

$$\bar{x} - z_{\frac{\alpha}{\text{۲}}} \cdot \frac{S}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{\text{۲}}} \cdot \frac{S}{\sqrt{n}}$$

$$\text{۳۶۴/۷} - z_{\text{۰/۰۵}} \cdot \frac{S}{\sqrt{n}} < \mu < \text{۳۶۴/۷} + z_{\text{۰/۰۵}} \cdot \frac{S}{\sqrt{n}}$$

$$\text{۳۶۴/۷} - \text{۱/۶۴۵} \times (\text{۱۲/۹۵}) < \mu < \text{۳۶۴/۷} + \text{۱/۶۴۵} \times (\text{۱۲/۹۵})$$

$$\text{۳۴۳/۴} < \mu < \text{۳۸۶}$$

مثال : پژوهشگری تأثیر روشی خاص را بر میانگین زمان انجام کاری مشخّص بررسی می‌کند. نمونه‌ی تصادفی از ۱۶ گروه این کار را بطور متوسّط در زمان ۲۵/۹ دقیقه با انحراف معیار ۳/۶ دقیقه انجام داده‌اند. فرض کنید زمان انجام کار از توزیعی پیروی می‌کند که تقریباً نرمال است. برای میانگین زمان انجام کار با این روش یک فاصله‌ی اطمینان ۹۵ درصدی درنظر گرفته و برآورد فاصله‌ای را تحلیل کنید.

$$1-\alpha = 95\% \longrightarrow \alpha = 0/05 \longrightarrow \frac{\alpha}{2} = 0/025$$

$$L < \mu < V$$

$$\bar{x} - t_{\frac{\alpha}{2}, n-1} \times \frac{S}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2}, n-1} \times \frac{S}{\sqrt{n}}$$

$$25/9 - t_{0/025, 15} \times \frac{S}{\sqrt{n}} < \mu < 25/9 + t_{0/025, 15} \times \frac{S}{\sqrt{n}}$$

$$25/9 - 2/131 \times \frac{3/6}{\sqrt{16}} < \mu < 25/9 + 2/131 \times \frac{3/6}{\sqrt{16}}$$

$$23/982 < \mu < 27/818$$

با ضریب اطمینان ۹۵٪، میانگین جامعه بین ۲۳/۹۸۲ و ۲۷/۸۱۸ می‌باشد.

مثال : تعداد زیادی شرکت عضو یک اتّحادیه‌ی صنفی‌اند. برای بررسی قانون جدید حدّاقل دستمزد برای اعضای اتّحادیه، می‌خواهیم میانگین تعداد کارکنان که در شرکت‌های عضو دستمزد ساعتی می‌گیرند را برآورد کنیم. برای این منظور نمونه‌ی تصادفی به اندازه‌ی ۲۲۵ از شرکت‌های عضو انتخاب کرده و تعداد کارکنان هر شرکت را ثبت می‌کنیم. میانگین و انحراف معیار نمونه بصورت زیر است :

$\bar{x} = ۸/۳۱$

$S = ۴/۸$

یک فاصله اطمینان ۹۹ درصدی برای میانگین جامعه بدست آورید :

$$1 - \alpha = ۹۹\% \longrightarrow \alpha = ۰/۰۱ \longrightarrow \frac{\alpha}{۲} = ۰/۰۰۵$$

$$L < \mu < V$$

$$\bar{x} - z_{\frac{\alpha}{۲}} \cdot \frac{S}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{۲}} \cdot \frac{S}{\sqrt{n}}$$

$$۸/۳۱ - z_{۰/۰۰۵} \cdot \frac{S}{\sqrt{n}} < \mu < ۸/۳۱ + z_{۰/۰۰۵} \cdot \frac{S}{\sqrt{n}}$$

$$۸/۳۱ - ۲/۵۷۶ \times \frac{۴/۸}{\sqrt{۲۲۵}} < \mu < ۸/۳۱ + ۲/۵۷۶ \times \frac{۴/۸}{\sqrt{۲۲۵}}$$

$$۷/۴۹ < \mu < ۹/۱۳$$

یک نشریه صنعت کامپیوتر می‌خواهد میانگین درآمد سهام عادی شرکت‌های کامپیوتری را در سال گذشته برآورد کند. نمونه تصادفی به اندازه $n = ۱۴$ شرکت از بین شرکت‌های کامپیوتری استخراج شده است که درآمد عادی این شرکت‌ها بصورت زیر است :

۸/۳۷ ، ۳۱/۳۸ ، ۶/۵۲- ، ۲۴/۸۷ ، ۸/۴۷- ، ۱۷/۷۸ ، ۳/۱۵- ، ۷/۲۰ ، ۱۵/۴۵ ، ۲۴/۸۳ ، ۰/۳۵ ، ۱۱/۲۸ ، ۲۱/۰۱ و ۲۶/۶۸

فرض کنید درآمد سهام عادی از توزیع غیرنرمال پیروی می‌کند که از نرمال بودن انحراف زیادی ندارد. برای میانگین درآمد سهام یک فاصله اطمینان ۹۰٪ ایجاد کنید.

$$\bar{x} = \frac{\sum_{i=۱}^{۱۴} x_i}{n} = \frac{۲۶/۶۸ + ۲۱/۰۱ + ۱۱/۲۸ + ۰/۳۵ + ۲۴/۸۳ + ۱۵/۴۵ + ۷/۲۰ + -۳/۱۵ + ۱۷/۷۸ + -۸/۴۷ + ۲۴/۸۷ + -۶/۵۲ + ۳۱/۳۸ + ۸/۳۷}{۱۴} = ۱۲/۲۲$$

$$S^{۲} = \frac{\sum_{i=۱}^{n} x_i^{۲} - \frac{\left(\sum_{i=۱}^{n} x_i\right)^{۲}}{n}}{n - ۱}$$

$$S^{۲} = \frac{۲۶/۶۸^{۲} + ۲۱/۰۱^{۲} + ۱۱/۲۸^{۲} + ۰/۳۵^{۲} + ۲۴/۸۳^{۲} + ۱۵/۴۵^{۲} + ۷/۲۰^{۲} + -۳/۱۵^{۲} + ۱۷/۷۸^{۲} + -۸/۴۷^{۲} + ۲۴/۸۷^{۲} + -۶/۵۲^{۲} + ۳۱/۳۸^{۲} + ۸/۳۷^{۲} - \frac{(۱۷۱/۰۶)^{۲}}{۱۴}}{۱۴ - ۱}$$

$$S^{۲} = \frac{۴۳۰۱/۲ - \frac{۲۹۲۶۲}{۱۴}}{۱۳}$$

$$S^{۲} = \frac{۴۳۰۱/۲ - ۲۰۹۰/۱}{۱۳} = ۱۷۰/۰۹ \quad \Rightarrow \quad S = ۱۳/۰۴$$

$\bar{x} = ۱۲/۲۲$

$S = ۱۳/۰۴$

$n = ۱۴$

$۱ - \alpha = ۹۰\% \longrightarrow \alpha = ۰/۱ \longrightarrow \frac{\alpha}{۲} = ۰/۰۵$

$$L < \mu < V$$

$$\bar{x} - t_{\frac{\alpha}{۲}, n-۱} \times \frac{S}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{۲}, n-۱} \times \frac{S}{\sqrt{n}}$$

$$۱۲/۲۲ - t_{۰/۰۵, ۱۳} \times \frac{S}{\sqrt{n}} < \mu < ۱۲/۲۲ + t_{۰/۰۵, ۱۳} \times \frac{S}{\sqrt{n}}$$

$$۱۲/۲۲ - ۱/۷۷۱ \times \frac{۱۳/۰۴}{\sqrt{۱۴}} < \mu < ۱۲/۲۲ + ۱/۷۷۱ \times \frac{۱۳/۰۴}{\sqrt{۱۴}}$$

$$۶/۰۴ < \mu < ۱۸/۳۹$$

برآورد نقطه‌ای نسبت جامعه :

اگر نمونه‌ی n تایی بطور تصادفی از یک جامعه انتخاب کنیم و X تعداد عناصر این نمونه باشد، که دارای مشخّصه‌ی مورد نظر است، نسبت نمونه‌ای یعنی $\bar{p} = \frac{X}{n}$ بهترین برآوردکننده p نسبت به سایر برآوردکننده‌هاست. میانگین و واریانس آماره‌ی p از رابطه‌ی زیر بدست می‌آید.

$$E(\bar{p}) = p$$

$$\sigma^{۲}_{\bar{p}} = \frac{p(۱ - p)}{n} = \frac{p.q}{n}$$

برآوردکننده‌ی $\sigma^{۲}$:

اگر نسبت جامعه p معلوم نباشد، $\sigma^{۲}$ هم نامعلوم است و بصورت زیر برآورد می‌شود :

$$S^{۲}_{\bar{p}} = \frac{\bar{p}(۱ - \bar{p})}{n} \Rightarrow S_{\bar{p}} = \sqrt{\frac{\bar{p}(۱ - \bar{p})}{n}}$$

خطای برآورد p برای نمونه‌های بزرگ :

این خطا بصورت $d = \left| \bar{p} - p \right|$ تعریف می‌شود و احتمال زیر برای مقدار خطا برقرار است :

$$P(d < z_{\frac{\alpha}{۲}} . S_{\bar{x}}) = ۱ - \alpha$$

و برآورد فاصله‌ای p برای نمونه‌های بزرگ برابر است با :

$$\bar{p} - z_{\frac{\alpha}{۲}} . \sqrt{\frac{\bar{p}(۱ - \bar{p})}{n}} < \mu < \bar{p} + z_{\frac{\alpha}{۲}} . \sqrt{\frac{\bar{p}(۱ - \bar{p})}{n}}$$

مثال : از کفش‌هایی که طیّ یک فرآیند مشخّص تولید شده‌اند، نمونه‌ای به اندازه‌ی ۴۰۰ جفت انتخاب شد که ۴۱ جفت آنها در رده‌ی معیوب قرار گرفتند. یک فاصله‌ی اطمینان ۹۵ درصدی برای نسبت کفش‌های معیوبی که تحت این فرآیند تولید شده‌اند، بدست آورید :

$$\bar{p} = \frac{۴۱}{۴۰۰} = ۰/۱۰۲۵$$

$$S_{\bar{p}} = \sqrt{\frac{۰/۱۰۲۵ \times (۱ - ۰/۱۰۲۵)}{۴۰۰}} = ۰/۰۱۵$$

$$۱ - \alpha = ۹۵\% \longrightarrow \alpha = ۰/۰۵ \longrightarrow \frac{\alpha}{۲} = ۰/۰۲۵$$

$$\bar{p} - z_{\frac{\alpha}{۲}} \cdot \sqrt{\frac{\bar{p}\,(۱ - \bar{p})}{n}} < \mu < \bar{p} + z_{\frac{\alpha}{۲}} \cdot \sqrt{\frac{\bar{p}\,(۱ - \bar{p})}{n}}$$

$$۰/۱۰۲۵ - ۱/۹۶ \times (۰/۰۱۵) < \mu < ۰/۱۰۲۵ + ۱/۹۶ \times (۰/۰۱۵)$$

$$۰/۰۷۳۱ < \mu < ۰/۱۳۱۹$$

مثال : یک اتّحادیه‌ی صنفی نمونه‌ی تصادفی از ۱۰۰ عضو را برای تعیین برآورد p نسبت به شرکت‌هایی که کارکنان نیمه وقت ندارند انتخاب می‌کند. اگر در این نمونه ۱۰ شرکت دارای کارکنان نیمه وقت باشند،

الف) مقدار p را برآورد کنید.

ب) یک فاصله‌ی اطمینان ۹۵ درصدی برای p بدست آورید.

$$\bar{p} = \frac{۹۰}{۱۰۰} = ۰/۹$$

$$S_{\bar{p}} = \sqrt{\frac{۰/۹ \times (۱ - ۰/۹)}{۱۰۰}} = ۰/۰۳$$

$$۱ - \alpha = ۹۵\% \longrightarrow \alpha = ۰/۰۵ \longrightarrow \frac{\alpha}{۲} = ۰/۰۲۵$$

$$\bar{p} - z_{\frac{\alpha}{۲}} \cdot \sqrt{\frac{\bar{p}\,(۱ - \bar{p})}{n}} < \mu < \bar{p} + z_{\frac{\alpha}{۲}} \cdot \sqrt{\frac{\bar{p}\,(۱ - \bar{p})}{n}}$$

$$۰/۹ - ۱/۹۶ \times (۰/۳) < \mu < ۰/۹ + ۱/۹۶ \times (۰/۳)$$

$$-۰/۳۱۸ < \mu < ۰/۸۵۸$$

Statistics and Probability

Dr . Hengameh Mohammadi Nejad

Collector : Amin Daneshi