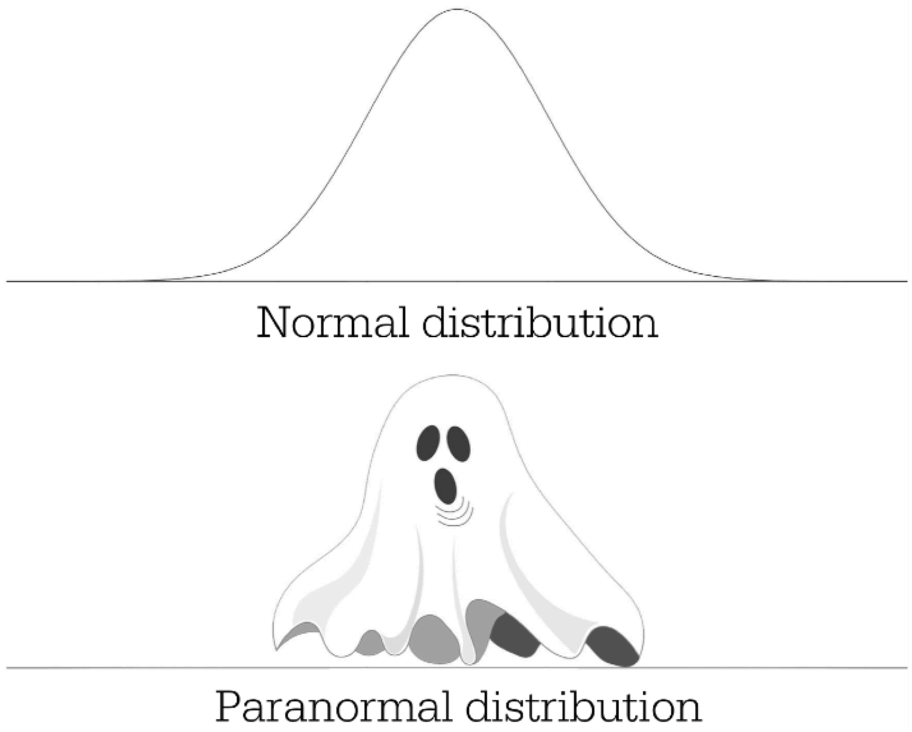
By:
Oleg Sidorenkov
Yaroslav Ulyanenkov
Pan ZhengWu

All tasks for Math Statistics test.



№1).

$E[X] = 1$
$E[Y] = 2$

$$E\left[2 \cdot X - Y + 2020\right] = 2 \cdot E[X] - E[Y] + 2020 = 2020$$

Ответ: A . 2020

№2).

$D[X] = 4$
$D[Y] = 8$
$k_{XY} = -1$

$$D\left[X - Y + 2021\right] = D[X - Y] = D[X] + D[Y] - 2 \cdot k_{XY} = 14$$

Ответ: A . 14

№3).

$D[X] = 4$
$D[Y] = 8$
$k_{XY} = -1$

$k_{X-2\cdot Y+2021;\ 3\cdot X-4\cdot Y-2022} = k_{X-2\cdot Y;\ 3\cdot X-4\cdot Y} = k_{X;\ 3\cdot X-4\cdot Y} - k_{2\cdot Y;\ 3\cdot X-4\cdot Y} = 3\cdot k_{XX} - 4\cdot k_{XY} - 6\cdot k_{XY} + 8\cdot k_{YY} = $
$=$
$= 3\cdot D[X] - 10\cdot k_{XY} + 8\cdot D[Y] = 86$

Ответ: A. 86

№4).

$D[X] = 4$
$D[Y] = 8$
$k_{XY} = -1$

$D\left[X - 2\cdot Y\right] = D[X] + 4\cdot D[Y] - 4\cdot k_{XY} = 40$

$D\left[3\cdot X - 4\cdot Y - 2022\right] = 9\cdot D[X] + 16\cdot D[Y] - 24\cdot k_{XY} = 188$

$$r_{X-2\cdot Y+2021;\ 3\cdot X-4\cdot Y-2022} = \frac{k_{X-2\cdot Y+2021;\ 3\cdot X-4\cdot Y-2022}}{\delta_{X-2\cdot Y+2021}\cdot\delta_{3\cdot X-4\cdot Y-2022}} = \frac{86}{\sqrt{188\cdot 40}} \approx 0.991721$$

Ответ: A. 0.991721

№5).

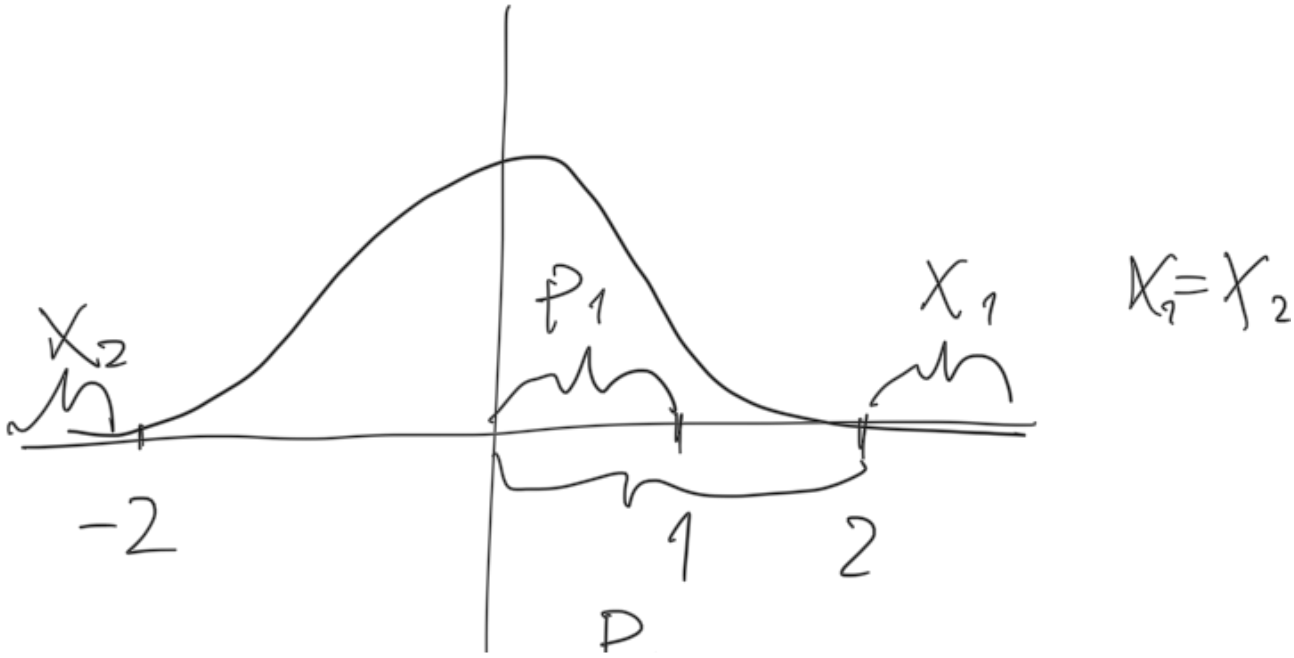
Пусть случайная величина $X \sim N(2,\ 4)$. Найдите $P\{-2 < X < 4\}$

$$P\{-2 < X < 4\} = P\{-4 < X - 2 < 2\} = P\left\{\frac{-4}{\sqrt{4}} < N(0{,}0) < \frac{2}{\sqrt{4}}\right\} = P\{-2 < N(0{,}0) < 1\}$$

Нормируем:

Найти квантиль 2 и 1 (P2 и P1) P2 + P1 = 0.3413 + 0.4773 = 0.8186

Ответ: A. 2020



№6).

$$P\,(X > c) = P\,(c < X < +\infty) = 0.1$$
$$P\,(c < X < +\infty) = 0.5 - \Phi_0\left(\frac{c-0}{1}\right) = 0.5 - \Phi_0\,(c) = 0.1$$
$$\Phi_0\,(c) = 0.4$$
$$c \approx 1.28$$

Ответ: $A\,.\ 1.28$

№7).

$$P\,(-c < X < c) = \Phi(c) - \Phi(-c) = 2\Phi(c) = 0.95$$

$$\frac{(1-0.95)}{2} = 0.025$$

$$\tau(0{,}025; 3) = -c \Rightarrow c = 3{,}1824$$

Ответ: $A\,.\ 3{,}1824$

№8).

$X^2(n) = \sum_{i=1}^{n} u_i^2$, где $u_i \sim N(0; 1)$

$$F(n, m) = \frac{\chi^2(n)/n}{\chi^2(m)/m}$$

$$\tau(n) = \frac{X}{\sqrt{\frac{\chi^2(n)}{n}}}$$

$$\frac{2X^2}{Y^2 + Z^2} = \frac{2\chi^2(1)/1}{\chi^2(2)} = \frac{\chi^2(1)/1}{\chi^2(2)/2} = F(1{,}2)$$

Ответ: A . $F(1,\ 2)$

№9).

$$P\left(X \le \sqrt{\frac{Y^2 + Z^2}{2}}\right) = ?$$

$$\sqrt{\frac{Y^2 + Z^2}{2}} = \sqrt{\frac{\chi^2\,(2)}{2}}$$

$$P\left(\frac{X}{\sqrt{\frac{\chi^2(2)}{2}}} \le 1\right) = P\left(\frac{X}{\sqrt{\frac{\chi^2(2)}{2}}} \le 1\right) = P\left(\tau\,(2) \le 1\right) \approx 0.7887$$

Ответ: A . 0.7887

№10).

$$V(X) = \begin{bmatrix} 4 & -1 & 0 \\ -1 & 3 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

$$D\left[X_1 - 2 \cdot X_2 - X_3\right] = D\left[X_1 - 2 \cdot X_2\right] + D\left[X_3\right] - 2 \cdot k_{X_1 - 2 \cdot X_2;\ X_3} =$$

$= D\left[X_1\right] + 4 \cdot D\left[X_2\right] - 4 \cdot k_{X_1X_2} + D\left[X_3\right] - 2 \cdot k_{X_1X_3} + 4 \cdot k_{X_2X_3} =$

$= 4 + 4 \cdot 3 - 4 \cdot (-1) + 2 - 2 \cdot 0 + 4 \cdot (-1) = 18$

Ответ: $A.\ 18$

№11).

$cov(X_1 + X_2,\ X_2 - X_3 + 1)$

$cov(X_1,\ X_2 - X_3 + 1) + cov(X_2,\ X_2 - X_3 + 1)$

$cov(X_1,\ X_2 - X_3) + cov(X_1,\ 1) + cov(X_2,\ X_2 - X_3) + cov(X_2,\ 1)$

$cov(X_1,\ X_2) - cov(X_1, X_3) + cov(X_2,\ X_2) - cov(X_2,\ X_3) = -1 - 0 + 3 + 1 = 3$

Ответ: $A.\ 3$

№12).

$Y_i = \beta \cdot x_i + \varepsilon_i$

Y_i	x_i
1	1
4	2
4	3

МНК-оценка β:

$$\widehat{\beta} = \frac{\sum_{i=1}^{n} x_i \cdot Y_i}{\sum_{i=1}^{n} x_i \cdot x_i} = 1.5$$

Ответ: $A.\ 1.5$

№13).

$Y_i = \beta \cdot x_i + \varepsilon_i$

Y_i	x_i
1	1
4	2
4	3

$$\widehat{\beta} = \frac{\sum_{i=1}^{n} x_i \cdot Y_i}{\sum_{i=1}^{n} x_i \cdot x_i} = 1.5$$

$$\widehat{Y} = X \cdot \widehat{\beta} = \begin{pmatrix} 1.5 \\ 3 \\ 4.5 \end{pmatrix}$$

Ответ: A . $\begin{pmatrix} 1.5 \\ 3 \\ 4.5 \end{pmatrix}$

№14).

14. Пусть задана регрессионная модель $Y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$. Имеются наблюдения:

Y_i	x_{i1}	x_{i2}
1	0	0
2	0	0
3	0	0
4	1	0
5	1	1

Чему равна сумма квадратов МНК-остатков RSS ?

A. 2 ;
B. 1;
C. 8;
D. 4;
E. Нет правильного ответа.

$$RSS = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$

$$\mathrm{RSS} = y^{\mathrm{T}} y - y^{\mathrm{T}} X (X^{\mathrm{T}} X)^{-1} X^{\mathrm{T}} y =$$

Где X матрица 3 * 5 (первая столбца является константой, всегда 1)

1	0	0
1	0	0
1	0	0
1	1	0
1	1	1

Ответ: A . 2

№15).

15. Пусть задана регрессионная модель $Y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$. Имеются наблюдения:

Y_i	x_{i1}	x_{i2}
1	0	0
2	0	0
3	0	0
4	1	0
5	1	1

Чему равен коэффициент детерминации R^2?

A. 0.8;
B. 0.9;
C. 0.2;
D. 0.6;
E. Нет правильного ответа.

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

$$TSS = \sum_{i=1}^{n} (Y_i - \bar{Y}_i)^2$$

$$ESS = \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y}_i)^2$$

Найти TSS & RSS:

$$TSS = (1-3)^2 + (2-3)^2 + (3-3)^2 + (4-3)^2 + (5-2)^2 = 10, RSS = 2$$

Ответ: A . 0.8

№16).

16. Пусть задана регрессионная модель $Y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$, причем $\varepsilon_1, \ldots, \varepsilon_n \sim \text{iid}(0, \sigma^2)$. Имеются наблюдения:

Y_i	x_{i1}	x_{i2}
1	0	0
2	0	0
3	0	0
4	1	0
5	1	1

Чему равна несмещённая оценка параметра σ^2?

A. 1;
B. 2;
C. 3;
D. 1.5;
E. Нет правильного ответа.

$$\hat{\sigma}^2 = \frac{RSS}{n-k-1} = \frac{2}{5-2-1} = 1$$

Ответ: $A.1$

№17).

17. Пусть задана регрессионная модель $Y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i$, причем $\varepsilon_1, \ldots, \varepsilon_n \sim \text{iid}(0, \sigma^2)$. Дана матрица

$$\hat{\mathrm{V}}(\hat{\beta}) = \begin{bmatrix} 1/3 & -1/3 & 0 \\ -1/3 & 4/3 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$

Чему равна оценка дисперсии $\hat{\mathrm{D}}(-\hat{\alpha} + \hat{\beta}_1 + 2\hat{\beta}_2)$?

A. $19/3$;
B. $17/3$;
C. $13/3$;
D. 3;
E. Нет правильного ответа.

$\hat{D}(-\alpha + \hat{\hat{\beta_1}} + 2\hat{\beta}_2) \to Dt\ = [-1,1,2]$
$Dt * \hat{V}(\hat{\beta}) * Dt^T\ = 6.333 = 19/3$
Ответ: $A.19/3$

№18).

18. Пусть задана модель линейной регрессии $Y_t = \beta \cdot t + \varepsilon_t$, $t = 1, 2$, в которой случайные ошибки удовлетворяют условиям $\mathbb{E}[\varepsilon_t] = 0$, $\mathrm{D}(\varepsilon_t) = \sigma^2$, $\mathrm{cov}(\varepsilon_s, \varepsilon_t) = 0$ при $s \neq t$. Рассматривается оценка неизвестного параметра β:

$$\tilde{\beta} = \frac{1 \cdot Y_1 + 2 \cdot Y_2}{1^2 + 2^2}.$$

Чему равно математическое ожидание оценки $\tilde{\beta}$?

A. β;
B. 2β;
C. $\frac{5}{3}\beta$;
D. $\frac{3}{5}\beta$;
E. Нет правильного ответа.

$E(Y) = E(\beta t + \varepsilon_i) = \beta t + E(\varepsilon) = \beta t$

$E(\tilde{\beta}) = E\left(\frac{1 * Y_1 + 2 * Y_2}{1^2 + 2^2}\right) = E\left(\frac{1 * \beta t_1 + 2 * \beta t_2}{1^2 + 2^2}\right) = E\left(\frac{1 * \beta + 4 * \beta}{1^2 + 2^2}\right) = E\left(\frac{5 * \beta}{5}\right) == E(\beta) = \beta$

Ответ: $A\ .\beta$

№19).

19. Пусть задана модель линейной регрессии $Y_t = \beta \cdot t + \varepsilon_t$, $t = 1, 2$, в которой случайные ошибки удовлетворяют условиям $\mathbb{E}[\varepsilon_t] = 0$, $D(\varepsilon_t) = \sigma^2$, $\text{cov}(\varepsilon_s, \varepsilon_t) = 0$ при $s \neq t$. Рассматривается оценка неизвестного параметра β:

$$\tilde{\beta} = \frac{1 \cdot Y_1 + 2 \cdot Y_2}{1^2 + 2^2}.$$

Чему равна дисперсия оценки $\tilde{\beta}$?

A. $\sigma^2 / 5$;
B. $\sigma^2 / 4$;
C. $\sigma^2 / 3$;
D. $3\sigma^2 / 5$;
E. Нет правильного ответа.

$$D(Y) = D(\beta t + \varepsilon_i) = D(\varepsilon) = \sigma^2$$

$$D(\tilde{\beta}) = D\left(\frac{1 * Y_1 + 2 * Y_2}{1^2 + 2^2}\right) = \frac{1}{25} * D\left(1 * Y_1 + 2 * Y_2\right) = \frac{1}{25} * (D(Y_1) + D(2Y_2)) = \frac{1}{25} * (\sigma^2 + 4 * \sigma^2) = \frac{1}{25} * (5 * \sigma^2) = \sigma^2 / 5$$

Ответ: $A . \sigma^2 / 5$

№20).

Ответ: A . 2020

№21).

Ответ: A . 2020

№22).

Ответ: A . 2020

№23).

Ответ: A . 2020

№24).

Ответ: A . 2020

№25).

Оценивается зависимость уровня заработной платы работника (wage) от уровня образования (educ), общего стажа (exper) и числа лет работы у текущего работодателя (tenure) в виде линейной регрессии $wage_i = \alpha + \beta_1 educ_i + \beta_2 exper_i + \beta_3 tenure_i + \varepsilon_i$,в которой ошибки $\varepsilon_1, \ldots, \varepsilon_n$ независимы и имеют нормальное распределение с нулевым математическим

ожиданием и дисперсией σ^2. Число наблюдений $n = 256$. В следующей таблице приведены результаты оценивания.

Переменная	Коэффициент	95% доверительный интервал
const	–2.87273	(–4.30480; –1.44067)
educ	0.598965	(0.498218; 0.699713)
exper	0.0223395	(–0.00134639; 0.0460254)
tenure	0.169269	(0.126747; 0.211790)

На уровне значимости 5% укажите, какие из переменных educ , exper и tenure оказывают значимое влияние на уровень заработной платы работника.

A. educ, tenure;
B. educ , exper ;
C. exper , tenure ;
D. educ , exper , tenure ;

E. Нет правильного ответа.

Доверительный интервал для коф β_j

$P\{T_1 \le \beta_j \le T_2\} = \gamma = P\{A \le T \le B\}$ где

$$A = tinv\left(\frac{1-\gamma}{2}, n-k-1\right) = -B$$

$$T = \frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}_{\hat{\beta}_j}}$$

Теперь для каждого аргумента $\beta_1, \beta_2, \beta_3, \alpha$ ищем $\hat{\sigma}_{\hat{\beta}_j}$ через доверительный интервал:

$$P(A \leq \frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}_{\hat{\beta}_j}} \leq B) = \gamma \to P(A\hat{\sigma}_{\hat{\beta}_j} \leq \hat{\beta}_j - \beta_j \leq B\hat{\sigma}_{\hat{\beta}_j}) = \gamma \to P(A\hat{\sigma}_{\hat{\beta}_j} - \hat{\beta}_j \leq -\beta_j \leq B\hat{\sigma}_{\hat{\beta}_j} - \hat{\beta}_j) = \gamma \to P(\hat{\beta}_j - A\hat{\sigma}_{\hat{\beta}_j} \geq \beta_j \geq \hat{\beta}_j - B\hat{\sigma}_{\hat{\beta}_j}) = \gamma$$

$$P(\hat{\beta}_j - B\hat{\sigma}_{\hat{\beta}_j} \leq \beta_j \leq \hat{\beta}_j - A\hat{\sigma}_{\hat{\beta}_j}) = \gamma$$

И $[\hat{\beta}_j - B\hat{\sigma}_{\hat{\beta}_j}, \hat{\beta}_j - A\hat{\sigma}_{\hat{\beta}_j}]$ это доверительный интервал который даны в условии

Сейчас ищем для edu (β_1) : (0.498218; 0.699713)

Левый границ это 0.498218

$$0.498218 = \hat{\beta}_1 - B\hat{\sigma}_{\hat{\beta}_1} \to \hat{\sigma}_{\hat{\beta}_1} = \frac{0.498218 - \hat{\beta}_1}{-B} = \frac{0.498218 - 0.598965}{-1.9643} = 0.051289$$

$\hat{\sigma}_{\hat{\beta}_j} = 0.051289$

Ставим в Т: $(T = \frac{\hat{\beta}_1}{\hat{\sigma}_{\hat{\beta}_1}} = \frac{0.598965}{0.051289} = 11.67824) > (B = 1.9645)$, где $\beta_1 = 0$ (гипотеза)

Так как T > B => не входит [A, B] => $\beta_1 \neq 0$ => edu (β_1) оказывают значимое влияние на уровень заработной платы работника на уровне значимости 5%

И так для exper и tenure

Ответ: $A\,.\; educ, tenure;$