Асимптотические доверительные интервалы (ДИ)

1. По выборке $X_1, \dots, X_n$,соответствующей распределению Пуассона $\Pi(\theta)$, постройте асимптотический доверительный интервал для параметра θ.

2. Для проверки качества деталей из большой партии выбрали 200 деталей. Среди выбранных деталей оказалось 12 бракованных. Постройте доверительный интервал уровня надёжности 0.95 для доли бракованных деталей.

3. Проводится две серии независимых испытаний по схеме Бернулли. В первой серии проведено n_1 испытаний, во второй - n_2 испытаний. Предполагается, что n_1 и n_2 достаточно велики. Построить асимптотический доверительный интервал для разности вероятностей "успеха" в двух сериях. (Показать, что статистика $\frac{\hat{p}_1-\hat{p}_2-(p_1-p_2)}{\sqrt{\widehat{D}(\hat{p}_1-\hat{p}_2)}}$, где $\widehat{D}(\hat{p}_1-\hat{p}_2)=\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}$, имеет асимптотически стандартное нормальное распределение).

4. Два года назад на втором курсе ОПИ ФКН обучалось 252 студента. На экзамене по ТВиМС 29 человек получили неудовлетворительные оценки. В прошлом учебном году на втором курсе ОПИ ФКН обучалось 286 студентов. На экзамене по ТВиМС 42 человека получили неудовлетворительные оценки. Постройте асимптотический доверительный интервал уровня надёжности 0.95 для разности вероятностей «успеха» по выборкам прошлого и позапрошлого года.

Проверка гипотез в биномиальной модели

5. Для проверки качества деталей из большой партии выбрали 200 деталей. Среди выбранных деталей оказалось 12 бракованных. Считается, что партия качественная, если она содержит не более 5% брака. Можно ли считать (на уровне значимости 0.05), что данная партия качественная? (! H_a односторонняя)

6. Известно, что вероятность рождения мальчика равна 0.52. В случайной выборке из 5000 человек в возрасте от 30 до 60 лет оказалось 2500 мужчин и 2500 женщин. Можно ли считать, основываясь на этих данных, что смертность среди мужчин и женщин одинакова? Уровень значимости считать равным 0.05. (! H_a односторонняя)

Домашнее задание

1. Из 500 опрошенных покупателей сети магазинов «Фиалка» 100 человек ответили, что они полностью довольны обслуживанием в данном магазине. Построить доверительный интервал уровня надёжности 0.95 для доли покупателей, довольных обслуживанием в сети «Фиалка».

2. Из 400 опрошенных покупателей сети «Жасмин» 70 человек ответили, что они довольны обслуживанием в данной сети. Постройте асимптотический доверительный интервал уровня надёжности 0.98 для разности долей покупателей, довольных обслуживание в сетевых магазинах «Фиалка» (см. предыдущую задачу) и «Жасмин».

3. Служба мониторинга опросила 1000 избирателей, среди которых оказалось 68 избирателей, собирающихся голосовать за партию А. Постройте 99% доверительный интервал для доли избирателей, собирающихся голосовать за партию А.

4. Для того, чтобы партия прошла в парламент, ей требуется набрать не менее 7% голосов избирателей. Служба мониторинга опросила 1000 избирателей, среди которых оказалось 68 избирателей, собирающихся голосовать за партию А. Можно ли считать (на уровне значимости 0.05), что партия А пройдёт в парламент? (! H_a односторонняя)

5. На экзамене по английскому языку студент должен ответить на 100 вопросов. На каждый вопрос предлагается 4 варианта ответа, один из которых правильный. Студент NN указал верные ответы на 30 вопросов. Можно ли считать (на уровне значимости 0.05), что студент NN не знает английский язык (т.е., выбор ответа на вопросы является случайным)? (! H_a односторонняя)

6. Известно, что женщины- водители составляют 30% от общего числа водителей. Из 635 зафиксированных ДТП 132 ДТП произошли по вине женщин-водителей. Можно ли считать (на уровне значимости 0.01), что женщины водят машину аккуратнее мужчин? (! H_a односторонняя)

$\hat\theta$ − сост. оценка θ

$\sqrt{n}(\hat\theta-\theta) \sim N(0, \sigma^2(\theta))$

по теор о наслед сход.

$\hat\theta \xrightarrow{P} \theta \Rightarrow \sigma(\hat\theta)\to\sigma(\theta)$

АДИ: $\left(\hat\theta \pm z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma(\hat\theta)}{\sqrt n}\right)$

1) $x_1 \dots x_n \sim \Pi(\theta) \Rightarrow EX=\theta$, $DX = \theta \Rightarrow \sigma=\sqrt\theta$, $\sigma^2_x = DX$

$\hat\theta=\bar x$ состоят и эффективная оценка θ

$\underbrace{\sigma(\hat\theta)}_{f(\hat\theta)} = \sqrt{\hat\theta}=\sqrt{\bar x}$

ДЗ: N2, N3 и N4 (нужно использовать N3)

$\hat p \to p \Rightarrow \sigma(\hat p)\to\sigma(p)$

$x_i \sim \Pi(\theta) \Rightarrow EX=\theta$, $DX=\theta$

$\hat\theta=\bar x$ − сост. и эффективн

$\bar x = \frac1n\sum_{i=1}^n x_i$

$E\bar x = E\hat\theta=\theta$ $\quad E\left(\frac1n\sum_{i=1}^n x_i\right)$

$D\bar x = \frac{Dx}{n}=\frac{\theta}{n}$

$\frac{\hat\theta - E\hat\theta}{\sqrt{D\hat\theta}} = \frac{\hat\theta-\theta}{\sqrt{\frac{\theta}{n}}} \overset{\text{ц.п.т}}{\sim} N(0,1)$

$P(T_1<\theta<T_2)=\gamma$

$\alpha = 1-\gamma$

$P\left(z_{\frac\alpha2} < \frac{\hat\theta-\theta}{\sqrt{\frac{\theta}{n}}} < z_{1-\frac\alpha2}\right) = \gamma$

$\hat\theta - z_{1-\frac\alpha2}\cdot\sqrt{\frac{\hat\theta}{n}} < \theta < \hat\theta + z_{1-\frac\alpha2}\cdot\sqrt{\frac{\hat\theta}{n}}$

$\theta \in \left(\bar x \pm z_{1-\frac\alpha2}\cdot\sqrt{\frac{\bar x}{n}}\right)$

$\hat\theta$ − сост и эффектив оценка θ, что при $n\to\infty$

$\sqrt n(\hat\theta-\theta)\sim N(0, \sigma^2(\theta) = f(\theta))$

$\hat\theta \xrightarrow{P}\theta \Rightarrow f(\hat\theta)\xrightarrow{P} f(\theta)$

$\sigma(\hat\theta) \xrightarrow{P}\sigma(\theta)$

Домашняя работа

2. Для проверки качества деталей из большой партии выбрали 200 деталей. Среди выбранных деталей оказалось 12 бракованных. Постройте доверительный интервал уровня надёжности 0.95 для доли бракованных деталей.

$n=200$, $x=12$, $\gamma=0{,}95$, $\alpha=1-0{,}95$

$\hat p$ − доля дефектных деталей в выборке

$\hat p = \frac{x}{n} = \frac{12}{200} = 0{,}06$

$1-\frac{\alpha}{2} = 1-0{,}025 = 0{,}975$

$z_{0{,}025} \approx 1{,}96$

$Se = \sqrt{\frac{\hat p(1-\hat p)}{n}} = \sqrt{\frac{0{,}06(1-0{,}06)}{200}} = \sqrt{0{,}000282} \approx 0{,}0168$

довери интервал:

нижняя граница

$0{,}06-1{,}96\cdot0{,}0168 = 0{,}06-0{,}0329 = 0{,}0271$

верхняя граница

$0{,}06+1{,}96\cdot0{,}0168 = 0{,}06+0{,}0329 = 0{,}0929$

Ответ: $\{0{,}027, 0{,}093\}$

известно, что АДИ для $\theta \left(\hat\theta - z_{1-\frac\alpha2}\cdot\frac{\sigma(\hat\theta)}{\sqrt n}, \hat\theta + z_{1-\frac\alpha2}\frac{\sigma(\hat\theta)}{\sqrt n}\right)$

$\hat\theta = \bar x \Rightarrow \sigma(\hat\theta) = \sqrt{\hat\theta} = \sqrt{\bar x}$

$\left(\bar x \pm z_{1-\frac\alpha2}\frac{\sqrt{\bar x}}{\sqrt n}\right)$

То что черным это я решала, забейте на черный шрифт

$S_n = \sum x_i \sim Bi(p,n)$

$x\sim Bi(p,1)$

$w_n = \frac{S_n}{n} = \frac{\sum x_i}{n} = \bar x$

по Th Муавра − Лапласа

$N\left(p, \frac{pq}{n}\right) \Rightarrow \frac{\overset{\bar x}{w_n} - p}{\sqrt{\frac{pq}{n}}} \sim N(0,1)$

$\Downarrow$

$\sigma = \sqrt{\frac{p(1-p)}{n}}$

$\hat p - p \Rightarrow \sigma(\hat p)\to\sigma(p) \Rightarrow \sigma(\hat p) = \sqrt{\frac{\hat p(1-\hat p)}{n}}$

$-z_{1-\frac\alpha2} \le \frac{w_n-p}{\sqrt{\frac{p(1-p)}{n}}} \le z_{1-\frac\alpha2}$

$\hat{p} = \overline{X}$ — сост и эффектив

$$-z_{1-\frac{\alpha}{2}} \le \frac{\overline{X}-p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \le z_{1-\frac{\alpha}{2}}$$

$$-z_{1-\frac{\alpha}{2}} \le \frac{\overline{X}-p}{\sqrt{\frac{\overline{X}(1-\overline{X})}{n}}} \le z_{1-\frac{\alpha}{2}}$$

$$-z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\overline{X}(1-\overline{X})}{n}} \le \overline{X}-p \le z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{\overline{X}(1-\overline{X})}{n}}$$

$$\hat{p} = \overline{X} = \frac{k}{n} = \frac{12}{200} = 0{,}06$$

$$0{,}06 - 1{,}96 \cdot \sqrt{\frac{0{,}06 \cdot 0{,}94}{200}} \le p \le 0{,}06$$

$$0{,}0271 \le p \le 0{,}0929$$

3. Проводится две серии независимых испытаний по схеме Бернулли. В первой серии проведено n_1 испытаний, во второй - n_2 испытаний. Предполагается, что n_1 и n_2 достаточно велики. Построить асимптотический доверительный интервал для разности вероятностей "успеха" в двух сериях. (Показать, что статистика $\frac{\hat{p}_1-\hat{p}_2-(p_1-p_2)}{\sqrt{\widehat{D}(\hat{p}_1-\hat{p}_2)}}$, где $\widehat{D}(\hat{p}_1-\hat{p}_2) = \frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}$, имеет асимптотически стандартное нормальное распределение).

$x_1, \dots, x_{n_1}$; $x_i \sim Bi(1, p_1)$ $\qquad \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \sim N\left(p_1; \frac{p_1 q_1}{n_1}\right)$; $\overline{x} \to p_1$

$y_1 \dots y_{n_2}$; $y_i \sim Bi(1, p_2)$ $\qquad \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \sim N\left(p_2; \frac{p_2 q_2}{n_2}\right)$ $\overline{y} \to p_2$

$$(\overline{x}-\overline{y}) \sim N(p_1-p_2; D(\overline{x}-\overline{y}))$$

$$D(\overline{x}-\overline{y}) = D(\overline{x}) + D(\overline{y}) - 2\operatorname{cov}(\overline{x};\overline{y}) \xlongequal{\overline{x},\overline{y}\text{ — независ}} D(\overline{x}) + D(\overline{y})$$

$$(\overline{x}-\overline{y}) \sim N(p_1-p_2, D(\overline{x}) + D(\overline{y}))$$

$$(\overline{x}-\overline{y}) \sim N\left(p_1-p_2; \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}\right)$$

$$\frac{(\overline{x}-\overline{y})-(p_1-p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \sim N(0,1)$$

по т. о наследовании необход $\hat{p} \xrightarrow{P} p_1 \Rightarrow \sigma(\hat{p}_1) \xrightarrow{P} \sigma(p_1)$

$$\frac{(\overline{x}-\overline{y})-(p_1-p_2)}{\underbrace{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}_{=\sigma}} \sim N(0,1)$$

$$P(T_1 \le p_1-p_2 \le T_2) = \gamma$$

$$-z_{1-\frac{\alpha}{2}} = z_{\frac{\alpha}{2}} \le \frac{\overline{x}-\overline{y}-(p_1-p_2)}{\sigma} \le z_{1-\frac{\alpha}{2}}$$