Вам предлагаются два типа задач: задачи 1-7 тестовые. Решать их следует, расставляя «да», «нет» или ответы на листке с заданием, рядом с пунктами. Не разрешается обводить варианты ответов кружками и использовать знаки $+$ и $-$ вместо «да» и «нет». Задачи 8-16 следует решать на отдельном листе бумаги, сохраняя нумерацию из задания. Переписывать условия задач не нужно!

1. (**1/2 балла**). Пусть θ^* — состоятельная оценка параметра θ. Выберите утверждения, которые следуют из этого факта:

а) θ^* — несмещенная оценка; **НЕТ**

б) θ^* — асимптотически несмещенная оценка; **НЕТ**

в) θ^* — асимпт. нормальная оценка; **НЕТ**

г) θ^*+5 — состоятельная оценка параметра $\theta+5$; **ДА**

д) $h(\theta^*)$ — состоятельная оценка параметра $h(\theta)$, если h – измеримая функция; **НЕТ**

е) $h(\theta^*)$ — состоятельная оценка параметра $h(\theta)$, если h – непрерывная функция; **ДА**

ё) $h(\theta^*)$ — состоятельная оценка параметра $h(\theta)$, если h – гладкая функция. **ДА**

б) Из сходимости по вероятности не следует сходимость математических ожиданий! Знакомый всем пример из т.в.: последовательность, сходящаяся к нулю по вероятности, с мат.ожиданиями стремящимися куда угодно, хоть к бесконечности — $\mathbf{P}(\xi_n=n^5)=1/n=1-\mathbf{P}(\xi_n=0)$. Требуются дополнительные условия типа равномерной интегрируемости, или попроще – существования интегрируемой мажоранты, ограниченности, и т.д.
ё) Почему-то многие отвечали «да» в пункте (е) и «нет» в пункте (ё)! Монетку, подбрасываемую наудачу, в таких случаях нужно контролировать :)

2. (**1/2 балла**). Пусть элементы выборки $X_1,\ldots,X_n$ имеют распределение с конечной дисперсией σ^2, $S^2=\overline{(X-\overline{X})^2}$, $S_0^2=\frac{1}{n-1}\sum_1^n(X_i-\overline{X})^2$. Выберите верные утверждения:

а) $\mathbf{D}\overline{X}=n\mathbf{D}X_1$; **НЕТ**

б) $S^2=\overline{X^2}-\left(\overline{X}\right)^2$; **ДА**

в) $S_0^2=\frac{n-1}{n}\sigma^2$; **НЕТ**

г) $\mathbf{E}S^2=\sigma^2$; **НЕТ**

д) $\mathbf{E}S^2\to\frac{n-1}{n}\sigma^2$; **НЕТ**

е) $\frac{(n-1)}{\sigma^2}S_0^2 \Subset \mathbf{H}_n$; **НЕТ**

ж) $\mathbf{E}S^2=\frac{n-1}{n}\sigma^2$; **ДА**

з) S_0^2 – состоятельная и несмещенная оценка σ^2. **ДА**

д) Предлагается оценить очевидную глупость — после перехода к пределу при $n\to\infty$ (очевидно), в пределе снова появляется n!
е) Еще одна глупость. При произвольном распределении выборки этот факт (даже если поставить $n-1$ степень свободы :)) места иметь не может! Попробуйте строго обосновать (используя наиболее простые аргументы), почему это не так, если элементы выборки имеют (дискретное!) распределение Бернулли или столь же дискретное распределение Пуассона, или равномерное распределение (и ограничены длиной отрезка).
з) S_0^2 – несмещенная выборочная дисперсия, не так ли?

3. (**1/3 балла**). Дана выборка объема n. Для проверки некоторой простой гипотезы H_1 против простой альтернативы используется критерий δ. Пусть $\alpha_i(\delta)$ — вероятность ошибки i-го рода. Верно ли, что:

а) всегда $\alpha_1(\delta)+\alpha_2(\delta)=1$? **НЕТ**

б) $\alpha_1(\delta)=\mathbf{P}_{H_1}(H_1 \text{ неверна})$? **НЕТ**

в) $\alpha_1(\delta)\to 0$ при $n\to\infty$? **НЕТ**

г) состоятельностью критерия называется предел $\lim_{n\to\infty}\alpha_2(\delta)$? **НЕТ**

д) δ — НМК? **НЕТ**

е) $\alpha_2(\delta)\to 0$? **НЕТ**

Абсурд ровно во всех пунктах.

б) Ни один критерий не решает вопрос, верна или не верна основная гипотеза. Он лишь принимает или отвергает ее. Приведенная вероятность равна нулю, и никакого отношения к критерию не имеет (сие – мое особое мнение. Н.Ч.). Пункт не оценивался из-за возможной неоднозначности терминологии.

4. (**1/3 балла**). Пусть с. в. χ_k^2 имеет распределение «хи-квадрат» с k степенями свободы. Найти слабый предел последовательности случайных величин $\dfrac{2-k}{\chi_k^2}$ при $k \to \infty$. Ответ: **-1**

ЗБЧ, определение распределения «хи-квадрат», наконец – способность распределения Стьюдента сходиться слабо к нормальному распределению (в доказательстве все это используется), в конце концов – слабая и по вероятности сходимости эквивалентны при сходимости к постоянной!

5. (**1/3 балла**). Пусть $\theta^* \in K_{2\theta}$ — оценка параметра θ.
Построить несмещенную оценку параметра θ. Ответ: $\theta^*/3$

Ответившим $\theta^* - 2\theta$ предлагается еще раз осознать: в задачах математической статистики любые ответы могут зависеть **только** от данных, или известных, объектов (выборки, например), но никак не от неизвестных (неизвестного параметра), и предназначаются как раз для получения разного рода информации об этом неизвестном параметре.

6. (**1/2 балла**). Выполнена ли первая часть условий регулярности (условие (R1)) для следующих семейств распределений, зависящих от параметра θ:

а) $\mathbf{U}(-\theta, 0)$, $\theta > 0$? **НЕТ**

б) с плотностью $f_\theta(y) = \theta e^{-\theta y} I(y > 0)$, $\theta > 0$? **ДА**

в) с плотностью $f_\theta(y) = e^{\theta + y} I(y < -\theta)$, $\theta > 0$? **НЕТ**

г) $\mathbf{B}(5, \theta)$, $0 < \theta < 1$? **ДА**

Пусть семейство $\mathbf{P}_\theta$ абсолютно непрерывно относительно некоторой меры μ, и имеет плотность $f_\theta(y) = \frac{d}{d\mu}\mathbf{P}_\theta((-\infty, y))$. Тогда оно удовлетворяет условию (R1), если для почти всех (относительно меры μ) точек y функция $\ln f_\theta(y)$ (как функция переменной θ) принадлежит классу $C^1(\Theta)$, то есть :-) непрерывна, дифференцируема и ее первая производная также непрерывна во всех точках $\theta \in \Theta$.

в) Рассмотрим даже не логарифм плотности, а саму плотность. При любом y плотность монотонно возрастает по θ при $\theta < -y$, и равна нулю при $\theta \geqslant -y$, то есть разрывна.

7. (**1/2 балла**). Является ли статистика T достаточной для неизвестного параметра θ семейства распределений P_θ, если:

а) $\theta = (a, \sigma^2)$, $P_\theta = \mathbf{N}(a, \sigma^2)$, $T = \overline{X^2}$? **НЕТ**

б) $\theta = \sigma^2$, $P_\theta = \mathbf{N}(3, \sigma^2)$, $T = \overline{X^2}$? **НЕТ**

в) $\theta = \sigma^2$, $P_\theta = \mathbf{N}(0, \sigma^2)$, $T = \overline{X^2}$? **ДА**

г) $\theta = (a, \sigma^2)$, $P_\theta = \mathbf{N}(a, \sigma^2)$, $T = (X_{(1)}, \ldots, X_{(n)})$? **ДА**

б) Функция правдоподобия представима в виде: $f_{\sigma^2}(\vec{X}) = (2\pi)^{-n/2}\ \sigma^{-n} \exp\left\{-\dfrac{1}{\sigma^2}\sum\limits_1^n (X_i - 3)^2\right\}$, и достаточной статистикой по факторизационной теореме Неймана-Фишера будет $\sum(X_i - 3)^2$, или (чуть длиннее) двумерная статистика $(\sum X_i, \sum X_i^2)$, но никак не удастся для $T = \frac{1}{n}\sum X_i^2$ представить ф-ю правдоподобия в виде $f_{\sigma^2}(\vec{X}) = h(\vec{X}) \cdot \Psi(T, \theta)$, где функции h и Ψ «зависят только от своих аргументов».

в) При известном математическом ожидании $T = \sum(X_i - a)^2$ — достаточная статистика для неизвестной дисперсии нормального распределения.

8. (**1/2 балла**). Пусть $F_n^*(y)$ — эмпирическая функция распределения, построенная по выборке объема n. Доказать, что при любом фиксированном y величина $F_n^*(y)$ — состоятельная и несмещенная оценка для $\mathbf{P}(X_1 < y)$.

См. соотв. утверждение для эмпирического распределения.

9. (**1/3 балла**). Пусть $X_1, \ldots, X_n$ — выборка из параметрического семейства распределений $\Pi(\lambda)$, где $0 < \lambda < \infty$. Является ли $2\overline{X}$ асимптотически нормальной оценкой? Если «да», то какого параметра и с каким коэффициентом? Если «нет» — почему?

АНО параметра 2λ, с коэфф. 4λ, — см. ЦПТ.

10. (**1/2 балла**). Пусть θ^* — асимптотически нормальная оценка параметра θ с коэффициентом $\sigma^2(\theta)$. Доказать, что $\dfrac{1}{(\theta^*)^2}$ — АНО для $\dfrac{1}{\theta^2}$. Найти коэффициент.

Для док-ва достаточно двух свойств: а) АНО состоятельна; б) слабо сходящуюся последовательность можно умножать на сходящуюся по вероятности к постоянной: если $\xi_n \Rightarrow \xi, \eta_n \xrightarrow{\text{p}} c$, то $\xi_n\eta_n \Rightarrow c\xi$. Коэффициент равен $4\sigma^2(\theta)/\theta^6$.

11. (**1 балл**). Производится $n \geqslant 2$ измерений неизвестного диаметра d круга. Предполагается, что измерения производятся с независимыми случайными ошибками, имеющими одинаковое нормальное распределение с нулевым средним и неизвестной дисперсией σ^2. Проверить несмещенность и состоятельность оценки $s^* = \dfrac{\pi}{4}\left(\left(\overline{X}\right)^2 - \dfrac{S_0^2}{n}\right)$ для площади круга, если $S_0^2 = (n-1)^{-1}\sum_1^n (X_i - \overline{X})^2$.

X_i — результаты измерений диаметра, имеют вид $X_i = d + ошибка_i$, и распределены нормально со средним d и неизвестной дисперсией σ^2. Так как S_0^2 — несмещенная выборочная дисперсия, имеем $\mathbf{E}s^* = \dfrac{\pi}{4}\left(\mathbf{D}\overline{X} + (\mathbf{E}\overline{X})^2 - \dfrac{\sigma^2}{n}\right) = \dfrac{\pi}{4}\left(\dfrac{\sigma^2}{n} + d^2 - \dfrac{\sigma^2}{n}\right)$ — площадь круга, то есть оценка несмещенная. Состоятельность еще более очевидна.

12. (**1/2 балла**). Сформулировать теорему Неймана-Пирсона для двух простых гипотез. Для какой постоянной C, участвующей в определении НМК, этот критерий совпадает с байесовским, если предположить, что априорные вероятности гипотез H_1 и H_2 равны, соответственно, $2/5$ и $3/5$?

13. (**2/3 балла**). По выборке объема 1 из нормального распределения с неизвестным средним a и неизвестной дисперсией построить минимаксный критерий для различения двух простых гипотез о параметре a.

Дана величина X_1 из распределения $\mathbf{N}(a,\sigma^2)$. Гипотезы: $a = a_1$ и $a = a_2$. Нарисовать на графике плотности, соответствующие каждой из гипотез и убедиться, что критерий, минимизирующий наибольшую из вероятностей ошибок 1-го и 2-го рода, устроен так: «$a = \min(a_1,a_2) \iff X_1 < (a_1+a_2)/2$». Для данных гипотез так же выглядит байесовский критерий при равных априорных вероятностях. Тот же ответ дает критерий отношения правдоподобия (из теоремы Неймана-Пирсона), в котором постоянная C ищется из условия равенства вероятностей ошибок $\alpha_1 = \alpha_2$ первого и второго рода.

14. Пусть $X_1, \ldots, X_n$ — выборка из параметрического семейства распределений $\mathbf{U}\left(\theta, \theta + \dfrac{\theta+2}{\theta+1}\right)$, где $\theta > 0$.

а) (**2/3 балла**). Найти ОМП для параметра θ.

б) (**1/3 балла**). Найти достаточную статистику для параметра θ.

Функция правдоподобия имеет вид: $f_\theta(\vec{X}) = \left(\frac{\theta+1}{\theta+2}\right)^n \mathbf{I}\left(\theta \leqslant X_{(1)}, X_{(n)} \leqslant \theta + \frac{\theta+2}{\theta+1}\right)$.

Достаточной статистикой является пара $(X_{(1)}, X_{(n)})$ (равно как и вся выборка или вариационный ряд :-)).

В области $\theta_0 \leqslant \theta \leqslant X_{(1)}$, где θ_0 — правый корень квадратного уравнения $X_{(n)} = \theta + \frac{\theta+2}{\theta+1}$ (если он есть), или в области $0 \leqslant \theta \leqslant X_{(1)}$ (если нет), функция правдоподобия монотонно возрастает по θ. Вне этой области она равна нулю. Так что ОМП — правая точка области $\theta_0 \leqslant \theta \leqslant X_{(1)}$. Ответ: ОМП $\hat{\theta}^* = X_{(1)}$.

15. (**1 балл**). Пусть $X_1, \ldots, X_n$ — выборка из нормального распределения $\mathbf{N}(a,1)$. Построить какой-либо критерий с ошибкой первого рода ε для различения гипотез H_1: $a = a_0$; H_2: $a < a_0$; H_3: $a > a_0$, «состоятельный» в следующем смысле: *вероятность принять гипотезу H_1, если она неверна, стремится к нулю с ростом n.*

Доказать, что построенный критерий действительно «состоятелен» в смысле данного определения.

Критерий строится с помощью функции $\sqrt{n}(\overline{X}-a_0)$, имеющей при верной первой гипотезе стандартное нормальное распределение. Если $\Phi_{0,1}(t_\varepsilon)=1-\varepsilon/2$, то критерий

$$\delta(\vec{X})=\begin{cases} H_2, \text{ если } \sqrt{n}(\overline{X}-a_0)<-t_\varepsilon; \\ H_1, \text{ если } -t_\varepsilon\leqslant\sqrt{n}(\overline{X}-a_0)\leqslant t_\varepsilon; \\ H_3, \text{ если } \sqrt{n}(\overline{X}-a_0)>t_\varepsilon \end{cases}$$

имеет вероятность ошибки первого рода ε и удовлетворяет условию:

$$\mathbf{P}_{\overline{H_1}}(\delta(\vec{X})\neq H_1)=\mathbf{P}_{\overline{H_1}}(\sqrt{n}|\overline{X}-a_0|>t_\varepsilon)\to 0,$$

поскольку $|\overline{X}-a_0|\to|a-a_0|>0$ п.н., где $a\neq a_0$ — истинное среднее, так что величина $\sqrt{n}|\overline{X}-a_0|$ с ростом n стремится к бесконечности (хоть п.н., хоть по вероятности).

16. (**1/2 балла**). Пусть $X_1,\ldots,X_n$ — выборка из распределения F с конечной и отличной от нуля дисперсией. Описать класс всех распределений F, для которых имеет место сходимость $\mathbf{P}(\overline{X}>3)\to\mathbf{P}(\mathbf{E}X_1>3)$.

Сначала — философия. Поскольку всегда при данных условиях конечен 1-й момент, то $\overline{X}\xrightarrow{\text{p}}\mathbf{E}X_1$. И поскольку $\mathbf{E}X_1$ – постоянная, эта сходимость эквивалентна слабой сходимости $\overline{X}\Rightarrow\mathbf{E}X_1$. Последняя эквивалентна поточечной сходимости функций распределения (или их хвостов) в точках непрерывности предельной функции распределения (или ее хвоста). Поэтому имеет место сходимость $\mathbf{P}(\overline{X}<\ (>)\ y)\to\mathbf{P}(\mathbf{E}X_1<\ (>)\ y)$ при всех $y\neq\mathbf{E}X_1$, но ее **может** не быть при $y=\mathbf{E}X_1$.
Теперь — решение. В данных условиях справедлива ЦПТ. Обозначим $\sigma^2=\mathbf{D}X_1$. Имеем

$$\mathbf{P}(\overline{X}>3)=\mathbf{P}\left(\frac{\sqrt{n}}{\sigma}(\overline{X}-\mathbf{E}X_1)<\frac{\sqrt{n}}{\sigma}(3-\mathbf{E}X_1)\right)\to\begin{cases} 0, \text{ если } 3-\mathbf{E}X_1<0; \\ 1/2, \text{ если } 3-\mathbf{E}X_1=0; \\ 1, \text{ если } 3-\mathbf{E}X_1>0. \end{cases}$$

В то же время $\mathbf{P}(\mathbf{E}X_1>3)$ равна нулю, если $\mathbf{E}X_1\leqslant 3$, или единице, если $\mathbf{E}X_1>3$, но никогда 1/2! То есть требуемый класс, как и ожидалось в философской части, совпадает с классом распределений, математическое ожидание которых отлично от 3.

А здесь могут быть Ваши комментарии:

Фамилия студента	Номер группы