Písemka 12. 6. 2015

Příklad 1. Hrajeme hru s náhodným počtem symetrických šestistěných kostek.Počet kostek K se řídí tímto rozdělením:

$$P[K=1]=(1-p)^2,\; P[K=2]=2p(1-p),\; P[K=3]=p^2,$$

kde $p \in (0,1)$. V každém hodu hodíme všemi K kostkami.

(a) Určete střední počet bodů padlých na kostkách v jednom hodu.
(b) Určete střední počet hodů, které musíme hodit předtím, než padne nejméně jedna trojka (padne-li trojka v prvním hodu je tedy tento počet 0).
(c) Pro $p = 1/2$ určete rozdělení počtu kostek za předpokladu, že součet na kostkách je 4.
(d)* Určete střední počet kostek za předpokladu, že součet na kostkách je 5.

Příklad 2. Předplatitelé (abonenti) koncertů filharmonie chodí na jednotlivé koncerty nezávisle na sobě a každý koncert navštíví s pravděpodobností $p \in (0,1)$.

(a) Nechť $p = 4/5$. Jakou kapacitu by musel mít sál, abychom s pravděpodobností alespoň 0,975 uspokojili všechny zájemce z řad předplatitelů o určitý koncert, prodáme-li 1200 předplatných?
(b) Nechť $p = 1/2$. Kapacita sálu je 450 míst. Chtěli bychom, aby se s pravděpodobností 0,95 daného koncertu mohl zúčastnit každý zájemce z řad předplatitelů. Kolik nejvýše můžeme prodat předplacených vstupenek?

Poznámka: V obou bodech (a) i (b) zdůvodněte postup řešení.

Příklad 3. Nechť $X_1, X_2, \ldots$ je posloupnost nezávislých náhodných veličin takových, že

$$\mathrm{P}(X_n = 0) = 1 - n^{-\alpha} \quad \text{a} \quad \mathrm{P}(X_n = \log n) = \frac{1}{n^\alpha}, \quad n = 1, 2, \ldots, \quad \alpha > 0$$

Dokažte, že

(a) $\overline{X}_n = n^{-1}\sum_{k=1}^n X_k$ konverguje pro $n \to \infty$ k nule skoro jistě.
(b) Dokažte, nebo vyvraťte tvrzení

$$P[X_n = 0,\ \text{pro nekonečně mnoho } n] = 1,\; P[X_n = \log n,\ \text{pro nekonečně mnoho } n] = 1$$

(c)* Dokažte, že X_n nekonverguje k nule skoro jistě pro $\alpha \in (0, \frac{1}{2}]$.

Příklad 4. Nechť $X_1, X_2, \ldots$ je náhodný výběr z rozdělení s hustotou

$$f(x) = \begin{cases} \exp(x+\mu), & x < -\mu, \\ 0, & x \geq -\mu, \end{cases}$$

kde $\mu \in \mathbb{R}$ je neznámý parametr.

(a) Najděte maximálně věrohodný odhad parametru μ.
(b) Diskutujte vlastnosti (nestrannost, konzistence) alternativního odhadu

$$\widetilde{\mu}_n = -\overline{X}_n - 1 = -\frac{1}{n}\sum_{i=1}^n X_i - 1.$$

Poznámky: Za každý příklad lze získat až 6 bodů. K úspěšnému napsání písemky je zapotřebí získat alespoň **12 bodů**. Tam, kde je hvězdička, lze získat dva body navíc.

Každou otázku pište, prosím, na **jiný papír**, na který nezapomeňte napsat své **jméno**!

Teoretická část 12. 6. 2015

Otázka 1. Definujte náhodný vektor, příslušnou distribuční funkci a marginální distribuční funkce. Napište vlastnosti sdružené distribuční funkce (stačí pro dvourozměrný náhodný vektor) a ukažte, že každá funkce s těmito vlastnostmi je sdružená distribuční funkce.

Nechť náhodný vektor (X, Y) má rozdělení se sdruženou hustotou

$$f(x,y) = \begin{cases} 2\,e^{-(x+2\,y)} & \text{pro } x \geq 0,\ y \geq 0, \\ 0 & \text{jinak.} \end{cases}$$

Určete marginální hustoty. Jsou X a Y nezávislé? Zdůvodněte odpověď.

Otázka 2. Zformulujte centrální limitní větu pro nezávislé náhodné veličiny.

Nechť $X_1, X_2, X_3, \ldots$ jsou nezávislé náhodné veličiny s Poissonovým rozdělením, kde $EX_i = \lambda_i > 0$. Zformulujte centrální limitní větu pro posloupnost $X_1, X_2, \ldots$.

Předpokládejte, že $\lambda_i = \lambda$ pro všechna i a využijte této znalosti ke konstrukci intervalového odhadu založeného na centrální limitní větě pro parametr λ.

Otázka 3. Vyslovte Raovu-Cramérovu nerovnost a vysvětlete její použití pro odhad parametru p v alternativním rozdělení

$$P[X = 1] = p = 1 - P[X = 0].$$

Co nám tato věta říká?

Otázka 4. Regresní model: popište matematicky model jednoduché lineární regrese. Ukažte, že odhady parametrů β_0 a β_1 metodou nejmenších čtverců

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \overline{x}_n) Y_i}{\sum_{i=1}^n (x_i - \overline{x}_n)^2}, \quad \widehat{\beta}_0 = \overline{Y}_n - \widehat{\beta}_1 \overline{x}_n$$

jsou nestranné. Odvoďte rozptyl odhadu $\widehat{\beta}_1$.

* Odvoďte rozptyl odhadu $\widehat{\beta}_0$.

Poznámky: Za každý příklad lze získat až 6 bodů. K úspěšnému napsání písemky je zapotřebí získat alespoň **12 bodů.** Tam, kde je hvězdička, lze získat dva body navíc.

Doporučujeme **nejdříve** každou otázku alespoň **stručně zodpovědět** (tj. např. zformulovat tvrzení, uvést definici, apod.) a až když Vám zůstane čas, tak se pouštět do podrobnější odpovědi (tj. např. důkazu tvrzení, odvozování, apod.).

Každou otázku pište, prosím, na **jiný papír**, na který nezapomeňte napsat své **jméno**!

Písemka 19. 6. 2015

Příklad 1. K výletu na kole je možné vyuzít kterýkoliv z n dní pobytu v hotelu. Celkově je v hotelu k hostů ($k \geq n$), kteří si zvolili náhodně a nezávisle na ostatních svůj den výletu. Určete pravděpodobnost, že každý den byl někdo z hostů na cyklovýletě.

* Jaká je pravděpodobnost, že každý den jeli na výlet alespoň dva hosté, je-li $k \geq 2n$?

Příklad 2. Pojišťovna pojišťuje 150 000 řidičů stejného automobilu. Každý řidič zaplatí 520 Kč ročního pojistného. S pravděpodobností 0,98 řidič během roku nezpůsobí nehodu. Pokud řidič způsobí nehodu, výše pojistného plnění (částky placené pojišťovnou) je 25 000 Kč.

(a) Na jaký zisk se může ředitel pojišťovny těšit na konci roku s pravděpodobností alespoň 0,95?

(b) S jakou pravděpodobností pojišťovna prodělá i svůj loňský zisk 200 000 Kč?

Řešte pomocí vhodné limitní věty, jejíž použití **zdůvodníte**.

Příklad 3. Automat na nápoje vydává kávu o udávaném objemu 150 ml. Vaším úkolem je ověřit tento údaj. Je-li střední hodnota náplně nižší než 152 ml, automat je nutné seřídit. Výběrový průměr objemu 25 náhodně vybraných káv je $\overline{X}_{25} = 150,8$ ml.

(a) Stačí vám tento údaj k tomu, abyste nechali stroj seřídit?

(b) Jak se rozhodnete, víte-li, že výběrový rozptyl je $S^2_{25} = 16$ ml^2?

Volte hladinu testu $\alpha = 0,05$. Doplňte předpoklady.

Příklad 4. Nechť $X_1, \ldots, X_n$ je náhodný výběr z rozdělení s konečným a nenulovým rozptylem.

(a) Dokažte, že pokud $\mathrm{E}\, X_1 = 0$, pak $\frac{1}{n}\sum_{i=1}^{n} X_i^2$ je nestranný a konzistentní odhad $\mathrm{var}(X_1)$.

(b) Dokažte, že odhad T_n definovaný jako

$$T_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2, \quad \text{kde } \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

je konzistentní odhad $\mathrm{var}(X_1)$.

(c)* Dokažte, že $\frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(X_i - X_j)^2$ je konzistentní odhad $\mathrm{var}(X_1)$.

(d)* Dokažte, $\frac{1}{2n}\sum_{i=2}^{n}(X_i - X_{i-1})^2$ je konzistentní odhad $\mathrm{var}(X_1)$.

Teoretická část 19. 6. 2015

Otázka 1. Definice nezávislosti obecného souboru náhodných veličin. Odvoďte rozdělení součtu dvou nezávislých nezáporných náhodných veličin s celočíselnými hodnotami.

Nechť X a Y jsou enzávislé stejně rozdělené náhodné veličiny s exponenciálním rozdělením, tj. s hustotou

$$f(x) = \begin{cases} \frac{1}{\mu}\exp(-x/\mu) & \text{je-li } x > 0, \\ 0 & \text{jinak}, \end{cases}$$

kde $\mu > 0$. Odvoďte rozdělení součtu $X + Y$.

Otázka 2. Vyslovte a dokažte Borelovu a Cantelliovu větu.

Uvažujte nezávislé stejně rozdělené náhodné veličiny $X_1, X_2, \ldots$ takové, že $\mathrm{E}\,|X_1| < \infty$. Pomocí Borelovy nebo pomocí Cantelliovy věty ukažte, že tato podmínka je ekvivalentní s podmínkou $\mathrm{P}[|X_n| > n \text{ pro nekonečně mnoho } n] = 0$

Otázka 3. Zformulujte centrální limitní větu. Odvoďte její důsledek pro Poissonovo rozdělení.

Nechť $X_1, \ldots, X_n$ je náhodný výběr z Poissonova rozdělení a $\overline{X}_n$ je výběrový průměr. Dokažte, že

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{\sqrt{n}(\overline{X}_n - \lambda)}{\sqrt{\overline{X}_n}} \le x\right) = \Phi(x), \quad \forall x \in \mathbb{R},$$

kde Φ je distribuční funkce normovaného normálního rozdělení. Využijte této znalosti ke konstrukci intervalového odhadu založeného na centrální limitní větě pro parametr λ.

Otázka 4. Bodový odhad parametru: vysvětlete úlohu odhadu, co je nestrannost a konzistence odhadu. Popište metodu maximální věrohodnosti.

Metodu maximální věrohodnosti a ověřování vlastností odhadů ilustrujte na následujícím příkladu: Uvažujte náhodný výběr z rozdělení s hustotou

$$f(x; a, \lambda) = \begin{cases} \exp\left(-\frac{x-a}{\lambda}\right) & \text{pokud } x > a \\ 0 & \text{jinak}, \end{cases}$$

kde $a > 0$ a $\lambda > 0$ jsou parametry.

(a) Najděte odhady $\widehat{a}$ a $\widehat{\lambda}$ parametrů a a λ metodou maximální věrohodnosti.

(b) Za předpokladu $a = 0$ najděte maximálně věrohodný odhad $\widetilde{\lambda}$ parametru λ.

(c) Jaké jsou vlastnosti odhadu $\widetilde{\lambda}$?

(d)* Jaké jsou vlastnosti odhadu $\widehat{a}$?

Poznámky: Za každý příklad lze získat až 6 bodů. K úspěšnému napsání písemky je zapotřebí získat alespoň **12 bodů**. Tam, kde je hvězdička, lze získat dva body navíc.

Doporučujeme **nejdříve** každou otázku alespoň **stručně zodpovědět** (tj. např. zformulovat tvrzení, uvést definici, apod.) a až když Vám zůstane čas, tak se pouštět do podrobnější odpovědi (tj. např. důkazu tvrzení, odvozování, apod.).

Každou otázku pište, prosím, na **jiný papír**, na který nezapomeňte napsat své **jméno**!

Písemka 1. 6. 2011

Příklad 1. Hodíme 400 krát symetrickou hrací kostkou. Označme Z_{400} součet dosažených výsledků. Určete přibližnou hodnotu pravděpodobnosti

$$P[1350 \leq Z_{400} \leq 1500].$$

Nezapomeňte svou aproximaci odůvodnit.

[5 bodů]

Příklad 2. Hostinský nakupuje od pivovaru sudové pivo, přičemž si vede evidenci, kolik piv z každého sudu vytočí—pivovar se dušuje, že v každém sudu je nejméně 100 piv. Po určité době dospěje k podezření, že pivovar jej okrádá a s pomocí svého známého statistika chce zjistit, zda tomu tak opravdu je. Hostinský chce, aby k falešnému obvinění pivovaru došlo s pravděpodobností nejvýše 0,01.

Statistik hostinskému poradí testovat na patřičné hladině spolehlivosti hypotézu, že v sudu je alespoň 100 piv. Jelikož to hostinský neumí, statistik za pár piv tento test provede sám. Jaká je rada hostinskému, víme-li se, že u 101 sudu vytočil hostinský následující počty piv:

sudů	15	25	21	25	15
piv v sudu	97	98	99	100	101

Uveďte také **všechny předpoklady**, za kterých statistik pracuje.
Uvědomte si, že výběrový rozptyl se nemění při přičtení libovolné konstanty.

[6 bodů]

Příklad 3. Dva hráči se střídají ve hře. Začíná hráč A, který při svém hodu vyhraje hru s pravděpodobností p, hráč B vyhraje při svém hodu s pravděpodobností q. Hra pokračuje, dokud jeden z hráčů nezvítězí.

(a) Uvažujte $p = q = 1/6$. S jakou pravděpodobností vyhraje hráč A a s jakou pravděpodobností hráč B?

(b) Uvažujte $p = q = 1/6$. Spočtěte pravděpodobnost, že počet všech hodů (hráče A spolu s hráčem B) je větší než dané k za podmínky, že vyhrál hráč B.

(c)* Jaké musí být p a q, aby pravděpodobnost výhry hráče B byla stejná jako pravděpodobnost výhry hráče A? (*toto nemusí mít jednoznačné řešení, navrhněte alespoň jedno*)

[6 bodů + prémie 2 body za (c)]

Příklad 4. Náhodné veličiny $X_1, \ldots, X_n$ jsou nezávislé a stejně rozdělené, X_i má hustotu

$$f(x) = (k-1)\theta^{k-1}x^{-k}, \quad \text{pro } x \geq \theta,$$

kde $\theta > 0$ je neznámý parametr, a $k > 4$ je známý parametr. Odvoďte odhad parametru θ metodou maximální věrohodnosti. Dokažte:

(a) Odhad

$$\widehat{\theta_n} = \overline{X_n}\frac{k-1}{k-2}$$

je nestranný a konzistentní.

(b)* Odhad

$$\widetilde{\theta_n} = \min_{1\leq i\leq n} X_i \frac{n(k-1)-1}{n(k-1)}$$

je nestranný a konzistentní.

[7 bodů + prémie 5 bodů za (b)*]

Poznámky: Na jedničku musíte spočítat všechny varianty. Na dvojku můžete vynechat příklad 4(b). Na trojku můžete navíc vynechat příklad 3(c).

K úspěšnému napsání písemky je zapotřebí získat alespoň **12 bodů**.

Otázky 1.6.2011

Otázka 1. Definice nezávislosti systému náhodných veličin.

(a) Podle čeho rozhodnete o nezávislosti pro konečný systém absolutně spojitých náhodných veličin a pro konečný systém diskrétních náhodných veličin?

(b) *Dokažte, nebo nejděte protipříklad:* Náhodné veličiny X, Y jsou nezávislé právě tehdy, když $\mathrm{E}XY = \mathrm{E}X\mathrm{E}Y$.

(c) *Dokažte, nebo najděte protipříklad:* Platí, že $[X \geq 0] \cap [Y \geq 0] = \emptyset$ právě tehdy, když X a Y jsou nezávislé.

(d) *Dokažte, nebo najděte protipříklad:* Jsou-li veličiny X a Y nezávislé, veličiny X a Z nezávislé a veličiny Y a Z nezávislé, pak i veličiny X, Y a Z jsou nezávislé.

Otázka 2. Bodový odhad parametru: vysvětlete úlohu odhadu, co je nestrannost a konzistence odhadu. Popište metodu maximální věrohodnosti.

Nad vaším domem prolétávají letadla jejichž rozestupy lze modelovat exponenciálním rozdělením s hustotou

$$f(x) = \frac{1}{\eta}\mathrm{e}^{-x/\eta}, \quad \text{pro } x > 0,$$

kde $\eta > 0$ je neznámý parametr.

Ve všední den je parametr rozdělení λ, o víkendech je parametr rozdělení $\lambda + \mu$. Předpokládejte, že ve všední dny jste napozorovali doby mezi přelety letadel $X_1, X_2, \ldots, X_n$ a o víkendu jste napozorovali doby $Y_1, Y_2, \ldots, Y_k$. Jak byste odhadli parametry λ a μ metodou maximální věrohodnosti? (stačí alespoň sestavit věrohodnostní rovnice a říci jak postupovat)

Otázka 3. Uvažujme posloupnosti $X_{n,j}$, $n = 1, 2, \ldots$, $j = 1, 2, \ldots, n$ nezávislých alternativních náhodných veličin s pravděpodobností úspěchu p_n. Jaké je limitní chování celkového počtu úspěchů $Y_n = \sum_{j=1}^{n} X_{n,j}$, pokud

(a) $n \to \infty$, $p_n \to 0$ a $np_n \to \mu \in (0, \infty)$.

(b) $n \to \infty$ a $p_n = p$ pro všechna n.

(c) $n \to \infty$, $p_n \to 0$ a $np_n \to 0$.

Svá tvrzení zdůvodněte.

Otázka 4. Testy hypotéz o střední hodnotě normálního rozdělení. Vysvětlete proč testová statistika a kritický obor mají daný tvar. Uvažujte regresní model

$$Y_i = \beta_1 + \beta_2 x_i + \sigma Z_i, \quad i = 1, \ldots, n,$$

kde x_i jsou známé konstanty (alespoň dvě z nich jsou různé), Z_i jsou nezávislé náhodné veličiny s normovaným normálním rozdělením a $\sigma > 0$ je neznámý parametr. Sestavte test hypotézy $H_0 : \beta_2 = 0$ proti alternativě $H_1 : \beta_2 \neq 0$. Co testuje tento test?

Poznámky: Pokud chcete jedničku, měli byste se pokusit o všechny varianty. Na dvojku můžete vynechat otázku 1(d) a otázku 3(c). Na trojku můžete dále vynechat otázku 1(c), v otázce 2 neuvažovat víkendový provoz a v otázce 4 uvažovat speciální dvouvýběrovou situaci, kdy $x_1 = \cdots = x_k = 0$ a $x_{k+1} = \cdots = x_n = 1$, kde $1 < k < n - 1$.