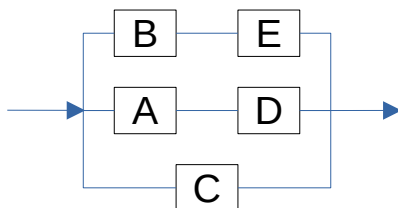


Pravděpodobnost a statistika - verze 220613B

Praktická část - (0 – 50) bodů, požadované minimum: 25 bodů

1. Systém se skládá z 5 komponent A-E (viz obrázek). Poruchy jednotlivých komponent jsou na sobě nezávislé. Pravděpodobnost poruchy jednotlivých komponent je následující: $P(A) = 1/7$, $P(B) = 1/4$, $P(C) = 1/2$, $P(D) = 1/6$, $P(E) = 1/4$.

- Určete pravděpodobnost poruchy bloků B + E a A + D. (5b)
- Určete pravděpodobnost poruchy celého systému. (4b)
- Určete pravděpodobnost, že celý systém bude v provozu. (1b)



2. Náhodný vektor $(X, Y)^T$ má sdruženou pravděpodobnostní funkci danou následující tabulkou:

$y \backslash x$	0	20	40
2	1/4	1/6	?
3	1/8	1/4	1/8

- Určete jaká je v tabulce chybějící hodnota sdružené pravděpodobnosti. (1b)
 - Určete a запиšte marginální pravděpodobnostní funkce náh. veličin X a Y . (2b)
 - Určete střední hodnotu a směrodatnou odchylku náhodných veličin X a Y . (2b)
 - Nechť $F(x; y)$ je sdružená distribuční funkce daného náhodného vektoru. Určete $F(25; 3)$. (1b)
 - Určete $P(X = 20 | Y = 3)$. (1b)
 - Určete korelační koeficient $\rho(X; Y)$ a na jeho základě posuďte míru lineární závislosti mezi náhodnými veličinami X a Y (3b)
3. Délku tresky obecné (v cm) lze modelovat normálním rozdělením se střední hodnotou 49,7 cm a směrodatnou odchylkou 3,6 cm.
- Načrtněte distribuční funkci délky tresky obecné. (1b)
 - Kolik procent tresek obecných měří nejvýše 53 cm? (2b)
 - Určete 5. percentil délky tresky obecné, interpretujte jej a zakreslete do náčrtku distribuční funkce z bodu a). (3b)
 - Je-li délka tresky obecné alespoň 53 cm, rybář ji na trhu označí jako „premium“ tresku a prodá ji dražší. Jaká je pravděpodobnost, že mezi 40 vylovenými rybami bude alespoň 5 premium tresek? (4b)

Popis datového souboru potřebného pro úlohy 4 a 5:

Přidělený datový soubor (R. A. Fisher (1936). "The use of multiple measurements in taxonomic problems". Annals of Eugenics. 7 (2): 179–188) obsahuje údaje o 50 rostlinách tří odrůd kosatců (Iris setosa, Iris virginica a Iris versicolor). U všech rostlin byly měřeny čtyři parametry: délka a šířka okvětních plátků (sepal) a okvětních lístků (petal) v centimetrech.

4. Okvětní plátky delší než 5,8 cm (*Sepal.Length* > 5,8) označme jako dlouhé. Analyzujte podíl kosatců Iris virginica a Iris versicolor s "dlouhými" okvětními plátky. Nezapomeňte na ověření předpokladů pro použití metod statistické indukce.
 - a) Pro kosatce Iris virginica a Iris setosa určete bodové a 90% oboustranné intervalové odhady pravděpodobnosti, že budou mít "dlouhý" okvětní lístek. Výsledek pro kosatce Iris virginica interpretujte. (3b)
 - b) Překračuje podíl kosatců Iris virginica s "dlouhými" okvětními plátky statisticky významně 50 %? K ověření využijte **vhodný** intervalový odhad i příslušný čistý test významnosti. (3b)
 - c) Na hladině významnosti 10 % určete, zda podíl kosatců Iris virginica s "dlouhými" okvětními lístky statisticky významně převyšuje podíl kosatců Iris versicolor s "dlouhými" okvětními lístky. Pro ověření použijte příslušný čistý test významnosti. (4b)
5. Na hladině významnosti 5 % rozhodněte, zda je statisticky významný rozdíl mezi průměrnými délkami okvětních plátků, popř. mediány délek okvětních plátků, jednotlivých druhů kosatců. V případě, že ano, nalezněte homogenní skupiny druhů kosatců a seřaďte je sestupně. **Obsahují-li data odlehlá pozorování, do analýzy je nezařazujte.** Nezapomeňte ověřit předpoklady pro použití zvoleného testu. **Poznámka:** V tomto příkladu do srovnání zařaďte všechny druhy kosatců (Iris setosa, Iris virginica a Iris versicolor). Délka okvětních plátků je v datovém souboru zaznamenána pod názvem Sepal.Length. (10b)