Poznámky z přednášek
Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

Pravděpodobnost a statistika

Peter Černo, 2010
petercerno@gmail.com

Garant: prof. RNDr. Jaromír Antoch, CSc.
E-mail: Jaromir.Antoch@mff.cuni.cz
Domácí stránka: `http://www.karlin.mff.cuni.cz/~antoch/`

Anotace: Zavedení základních pojmů a metod teorie pravděpodobnosti a matematické statistiky a příklady jejich aplikací. Jedná se zejména o pojem pravděpodobnosti, náhodné veličiny a jejího rozdělení, nezávislosti, náhodného výběru a jeho popisných charakteristik, konstrukci odhadů, testování hypotéz, náhodné generátory. Důraz je kladen na praktické použití metod s využitím dostupného statistického software.

Sylabus:

1. Základní pojmy teorie pravděpodobnosti - náhodné jevy, pravděpodobnost, podmíněná pravděpodobnost, věta o úplné pravděpodobnosti a Bayesova věta, nezávislost náhodných jevů

2. Náhodné veličiny a jejich rozdělení - náhodná veličina, diskrétní náhodná veličina, charakteristiky diskrétní náhodné veličiny, některé typy diskrétně rozdělených náhodných veličin, spojitá náhodná veličina, charakteristiky spojité náhodné veličiny, některé typy spojitě rozdělených náhodných veličin, centrální limitní věta

3. Náhodné vektory a jejich rozdělení - náhodný vektor, charakteristiky rozdělení náhodného vektoru, nezávislost náhodných vektorů, charakteristiky lineární kombinace náhodných veličin, vícerozměrné normální rozdělení

4. Úvod do matematické statistiky - náhodný výběr, uspořádaný výběr, přehled běžně užívaných popisných statistik

5. Teorie odhadu - bodové odhady, bodové odhady parametrů pro vybraná rozdělení, intervaly spolehlivosti

6. Teorie testování hypotéz - úvod do testování hypotéz, jedno- a dvouvýběrová analýza pro normální rozdělení, párový test, testy nulovosti korelačního koeficientu, test chí-kvadrát dobré shody

7. Regrese - lineární regrese s jednou vysvětlující proměnnou, lineární regrese s více vysvětlujícími proměnnými

8. Simulace - generátory náhodných čísel a základy simulací Monte Carlo

Cíl předmětu: Studenti se seznámí se základy teorie pravděpodobnosti a matematické statistiky. Hlavním cílem je pochopení podstaty statistických a pravděpodobnostních postupů prezentovaných v dalších přednáškách.

Literatura:

1. Anděl J., Statistické metody, MATFYZPRESS, Praha 1998.
2. Bartoszynski R. and Niewiadomska-Budaj M., Probability and Statistical Inference, J. Wiley, 1996.
3. Jarušková D., Matematická statistika, skriptum ČVUT, Praha 2000.
4. Zvára K. a Štěpán J., Pravděpodobnost a matematická statistika, MATFYZPRESS, Praha 1997.

This page is intentionally left blank.

02.10.2007 Tu
PETER ČERNO

MAI059 EX: PRAVDEPODOBNOST A ŠTATISTIKA

komarek@karlin.mff.cuni.cz
www.karlin.mff.cuni.cz/~komarek/vyuka.html

100b MAXIMUM:
- 20b - AKTÍVNA ÚČASŤ
- 40b - PÍSOMKA (NOVEMBER/DECEMBER)
- 40b - ZÁPOČTOVÁ ÚLOHA NA DOMA (PRED VIANOCAMI)
  www.r-project.org - PROGRAM R DO 29.02.2008

+ OPRAVNÁ PÍSOMKA (1. TÝŽDEŇ S.O.)

JE MOŽNÉ SPRAVIŤ SKÚŠKU BEZ ZÁPOČTU

NA ZÁPOČET .. 65 b !!

$(\Omega, \mathcal{A}, P)$

Ω – MNOŽINA MOŽNÝCH VÝSLEDKOV

$\mathcal{A}$ – ALGEBRA MNOŽIN NA Ω

$P: \mathcal{A} \to \langle 0,1 \rangle$

$P(A) = \frac{|A|}{|\Omega|}$

PRÍKLAD: 5 GÚľ (B, M, Č, Z, F)

3 GUĽE, AKÁ JE PRAVDEPODOBNOSŤ ŽE MEDZI NIMI JE
a) BIELA GUĽA
b) NIE JE MODRÁ GUĽA
c) JE BIELA A NIE JE MODRÁ

$\Omega = \binom{\{B, M, Č, Z, F\}}{3}$ $\quad |\Omega| = \binom{5}{3} = 10$

a) $|A| = \frac{1 \cdot 4 \cdot 3}{1 \cdot 2 \cdot 3} = 2$

b) $|A| = \frac{4 \cdot 3 \cdot 2}{1 \cdot 2 \cdot 3} = 4$

c) $|A| = \frac{1 \cdot 3 \cdot 2}{1 \cdot 2 \cdot 3} = 1$

a) AKÁ JE PRAVDEPODOBNOSŤ, ŽE PERMUTÁCIA $\pi \in S_n$ NEMÁ PEVNÝ BOD?

$A_i = \{\pi \in S_n \mid \pi(i) = i\}$

$$|A_1 \cup \ldots \cup A_n| = |A_1| + \ldots + |A_n| - |A_1 \cap A_2| - \ldots |A_{n-1} \cap A_n| + \ldots + (-1)^{n-1} |A_1 \cap \ldots A_n|$$

$$= \binom{n}{1}(n-1)! - \binom{n}{2}(n-2)! + \binom{n}{3}(n-3)! - \ldots + (-1)^{n-1}\binom{n}{n}(n-n)! =$$

$$= n! - \frac{n!}{2!} + \frac{n!}{3!} - \ldots + (-1)^{n-1}\frac{n!}{n!} =$$

$$= n!\left(\frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \ldots + (-1)^{n-1}\frac{1}{n!}\right)$$

$$p = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \ldots + (-1)^n \frac{1}{n!} \longrightarrow \frac{1}{e}$$

!! AKÁ JE PRAVDEPODOBNOSŤ, ŽE ASPOŇ PRÁVE k HOSTÍ DOSTANE KABÁT

n - ŠTUDENTOV

PRAVDEPODOBNOSŤ, ŽE 2 MAJÚ NARODENINY V ROVNAKÝ DEŇ

NÁHODNÁ VELIČINA $X : \Omega \to \mathbb{R}$

$$P(X = x) = P(\omega \in \Omega : X(\omega) = x)$$

$$11 = \overset{6}{1+4+6} = \overset{3}{1+5+5} = \overset{6}{2+3+6} = \overset{6}{2+4+5} = \overset{3}{3+3+5} = \overset{3}{3+4+4}$$

$$p = \frac{6+3+6+6+3+3}{6^3} = \frac{27}{216} = \frac{3^3}{6^3} = \underline{\underline{\frac{1}{8}}}$$

2/2

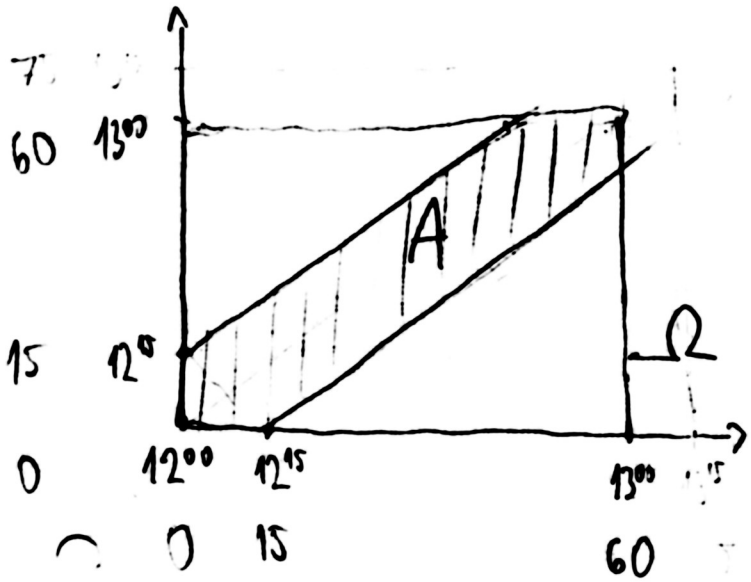
02.10.2007 Th
PETER ČERNO

MAI059 EX: PRAVDEPODOBNOSŤ A ŠTATISTIKA

Ω .. (NE)SPOČETNÁ MNOŽINA
$\mathcal{A}$ .. σ-ALGEBRA
$P : \mathcal{A} \to \langle 0,1 \rangle$

$P(\emptyset) = 0$, $P(\Omega) = 1$, $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$.

2 ĽUDIA MEDZI 12 A 13 HODINOU PRÍDU KU KOŇOVI A ČAKAJÚ 15 MIN.



$$P = 1 - \frac{45^2}{60^2} = 1 - \left(\frac{3}{4}\right)^2$$

100 OČÍSLOVANÝCH LÍSTKOV, VYBERIEM 3, MOŽNO ICH USPORIADAŤ DO ARITM. POSTUPNOSTI

$|\Omega| = \binom{100}{3}$

$a \in \{1, \dots, 100\}$

$|A| = a, a+d, a+2d$ $\quad d \in \mathbb{N}$, $a + 2d \leq 100$

$a \leq 100 - 2d$

$d = 1 .. 49$

$(100-2) + (100-4) + \dots + (100 - 2 \cdot 49) =$
$= 49 \cdot 100 - 2 \cdot (1 + \dots + 49) = 4900 - 2 \cdot 49 \cdot 25 = 49 \cdot 50 = 2450$

08.10.2007 Mo
PETER ČERNO

EX:

MAIOSY PRAVDEPODOBNOST A STATISTIKA

$(\Omega, \mathcal{A}, P)$

JAV < ELEMENTÁRNY / OBECNÝ

$0 \leq P(A_i) \leq 1 \qquad A_i \subseteq \Omega$

$P(\Omega) = 1$

$A_i \cap A_j = \emptyset$ SÚ NEZLÚČITEĽNÉ

$P(A) = \sum_{\omega_i \in A} P(\omega_i)$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

POJEM PODMIENENEJ PRAVDEPODOBNOSTI

BAYESOVA VETA (VETA O INVERZNEJ PRAVDEPODOBNOSTI)

DEF.: $P(A|B) \overset{def}{:=} \dfrac{P(A \cap B)}{P(B)}$ PRE $P(B) > 0$

PODMIENKA ZNAMENÁ ZÚŽENIE PRIESTORU VÝSLEDKOV NA MNOŽINU B A IBA MEDZI NIMI POČÍTAM TIE, KDE NASTALO A.

VETA: $P(A \cap B) = P(A|B) \cdot P(B)$

PRÍKLAD (LAPLACE, PASCAL 1793)

2 HRÁČI, HÁDZU MINCE

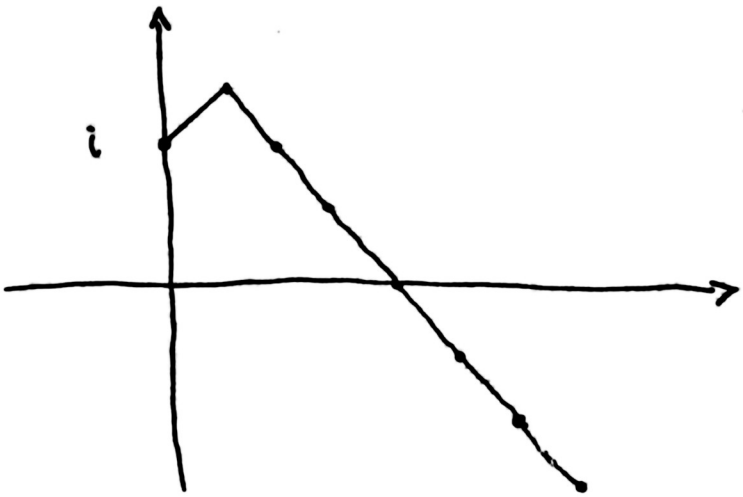
A .. i \$, B .. n-i \$.

$X_i \begin{cases} +1 & 1/2 \\ -1 & 1/2 \end{cases}$ POPIS JEDNEJ HRY

ZAUJÍMA NÁS $S_k = \left(\sum_{i=1}^{k} X_i\right) + S_0, \quad S_0 = i$

S_k POPISUJE STAV PEŇAŽENKY PO k HRÁČOCH, KEĎ SOM ZAČAL S i \$.

TZV. MODEL NÁHODNEJ PRECHÁDZKY.



STAVY PRECHÁDZKY SÚ $\{0, 1, \dots, n\}$

ZAČNEME V STAVE i,

$p_i = P(\text{VŠETKO VYHRÁM}) = P(A_i)$

<u>VETA O ÚPLNEJ PRAVDEPODOBNOSTI</u>: $(\Omega, \mathcal{A}, P)$

$\Omega = \bigcup H_i, \quad H_i \cap H_j = \emptyset \quad i \neq j$

$P(A) = P(A \cap \Omega) = P(A \cap (\bigcup H_i)) =$

$P(\bigcup A \cap H_i) = \sum P(A \cap H_i).$

$p_i = P(A_i | 1_+) P(1_+) + P(A_i | 1_-) P(1_-) =$

$= \frac{1}{2} p_{i+1} + \frac{1}{2} p_{i-1}$

VYJDE $p_i = \frac{i}{n}$

$\sum p_i = \frac{n+1}{2}$

08.10.2007
PETER ČERNO

EX: MA1059 PRAVDEPODOBNOST A ŠTATISTIKA

P(MÁ SKUTOČNE AIDS | TEST DOPADOL POZITÍVNE)

$$\left.\begin{array}{l} P(T+|A+) \\ P(T-|A-) \end{array}\right\}$$ TÁTO INFORMÁCIA BY MALA BYŤ ZNÁMA

$$P(A+|T+) = \frac{P(T+|A+)\,P(A+)}{P(T+)}$$

$P(A+)$ JE DANÉ

$$= \frac{P(T+|A+)P(A+)}{P(T+|A+)P(A+) + P(T+|A-)P(A-)}$$

BAYESOVA VETA: $(\Omega, \mathcal{A}, P)$, $\Omega = \bigcup H_i$, $H_i \cap H_j = \emptyset$ $i \neq j$

$A \subseteq \Omega$ $\quad P(H_i|A) = \frac{P(A|H_i)\,P(H_i)}{P(A)} =$

$$= \frac{P(A|H_i)\,P(H_i)}{\sum_j P(A|H_j)\,P(H_j)}$$

09.10.2007 TU
PETER ČERNO

EX: MA1059 PRAVDEPODOBNOSŤ A ŠTATISTIKA

URNA: M ROZLÍŠITEĽNÝCH GUĽÍ

m GUĽÍ ŤAHÁME

1) ZÁLEŽÍ / NEZÁLEŽÍ NÁS POREDIE

2) VRACIAME / NEVRACIAME GUĽE DO URNY

ZÁLEŽÍ NÁS POREDIE + VRACIAME DO URNY:

$$|\Omega| = M^m$$

ZÁLEŽÍ NÁS POREDIE + NEVRACIAME DO URNY:

$$|\Omega| = M \cdot (M-1) \dots (M-n+1)$$

NEZÁLEŽÍ NÁS POREDIE + VRACIAME DO URNY:

$$|\Omega| = \binom{M+m-1}{m}$$

o o ... o m GUĽIČIEK

_ _ _ .... _ M PRIEHRADOK

NEZÁLEŽÍ NÁS POREDIE + NEVRACIAME DO URNY:

$$|\Omega| = \binom{M}{m}$$

KOĽKO JE VÝSLEDKOV, KTORÉ VRACAJÚ 11

$$(x+x^2+\dots+x^6)^2 = x^2(1+x^2+x^3+x^4+x^5)^2 =$$

$$= x^2 \cdot \frac{1}{(1-x)^2}(1-x^6)^2 = x^2 \cdot \frac{1}{(1-x)^2} \cdot (1-2x^6+x^{12})$$

$$\frac{1}{1-x} = 1+x+x^2+\dots$$

$$\left(\frac{1}{1-x}\right)^2 = 1+2x+3x^2+4x^3+\dots$$

$$= (1+2x+3x^2+4x^3+5x^4+6x^5+7x^6+8x^7+9x^8+10x^9+\dots) \cdot \dots (x^2-2x^8+x^{14}) \Rightarrow$$

$$[x^{11}] = 10 - 2 \cdot 4 = \underline{\underline{2}}$$

$$(x+x^2+\dots+x^6)^3 = x^3 \cdot \frac{1}{(1-x)^3} \cdot (1-x^6)^3 =$$

$$= x^3 \cdot \frac{1}{(1-x)^3} \cdot (1-3x^6+3x^{12}-x^{18})$$

$$\frac{1}{1-x} = 1+x+x^2+x^3+x^4+\dots$$

$$\left(\frac{1}{1-x}\right)^2 = 1+2x+3x^2+4x^3+5x^4+\dots+(n+1)x^n+\dots$$

$$\left(\frac{1}{1-x}\right)^3 = 1+3x+(1+2+3)x^2+(1+2+3+4)x^3+\dots+\boxed{\frac{(n+1)(n+2)}{2}x^n}+\dots$$

$$= x^3\left(\dots+\frac{(2+1)(2+2)}{2}x^2+\dots+\frac{(9+1)(9+2)}{2}x^9+\dots\right)(1-3x^6+\dots)$$

$$[x^{11}] = -3\cdot\frac{3\cdot 4}{2}+\frac{10\cdot 11}{2} = -18+55 = 37$$

PRÍKLAD: MÁME r PREDMETOV, n PRIEHRADOK

PREDMETY $\begin{cases}\text{ROZLÍŠITEĽNÉ}\\ \text{NEROZLÍŠITEĽNÉ}\end{cases}$

MAXWELLOV - BOLTZMANOV MODEL

$S = (s_1, \dots, s_r) \qquad s_i \in \{1, \dots, n\}$

VŠETKY STAVY SÚ ROVNAKO PRAVDEPODOBNÉ.

KOĽKO JE MOŽNÝCH STAVOV? n^r

k_i = POČET OBJEKTOV V i-TEJ PRIEHRADKE

$P(k_i = k) = ?$

TJ. PRAVDEPODOBNOSŤ, ŽE NA k MIESTACH JE i ... !!

$$\left[\binom{r}{k}\cdot(n-1)^{r-k}\right] / n^r$$

09.10.2007 Tu
PETER ČERNO

EX: MAI059 PRAVDEPODOBNOST A STAT.

$$\lim \left(1+\frac{\frac{r_n}{n}}{r_n}\right)^{r_n} = e^{\lambda}$$

$$\lim \frac{r_n}{n} = \lambda$$

$r_n \to$ POČET OBJEKTOV ZÁVISÍ NA n

$r_n \to \infty, \quad \frac{r_n}{n} \to \lambda$

$$\lim_{n\to\infty} \frac{\binom{r_n}{k}(n-1)^{r_n-k}}{n^{r_n}} = \frac{r_n!}{(r_n-k)!\,k!}\cdot\left(1-\frac{1}{n}\right)^{r_n}\frac{1}{(n-1)^k} =$$

$$= \underbrace{\frac{r_n}{(n-1)}}_{\lambda}\frac{(r_n-1)}{(n-1)}\cdots\underbrace{\frac{(r_n-k+1)}{(n-1)}}_{\lambda}\left(1-\frac{1}{n}\right)^{r_n}\frac{1}{k!}$$

$$= \lambda^k\cdot\frac{1}{k!}\lim_{n\to\infty}\left(1-\frac{\frac{r_n}{n}}{r_n}\right)^{r_n} = \frac{\lambda^k}{k!}\cdot e^{-\lambda}$$

POISSONOVE ROZDELENIE

AKÁ JE PRIBLIŽNE PRAVDEPODOBNOST, ŽE NA MFF 3. ROČ 101 NEMÁ NIKTO NARODENINY 28.03.

$n=365, \quad r_n = 200, \quad \lambda = \frac{200}{365}$

$$P(k_{28.03}=0) \approx \frac{\lambda^0}{0!}e^{-\lambda} = e^{-\frac{200}{365}}$$

DOMÁCA ÚLOHA: SPOČÍTAŤ $E K_i$ = PRIEMERNÝ POČET PREDMETOV V 1 PRIEHRADKE

$$\sum_{k=0}^{r} k\, P(K_i = k)$$

BOSE-EINSTEINOV MODEL (PREDMETY SÚ NEROZLÍŠITEĽNÉ)

$r = (r_1, \dots, r_n)$ $\quad$ r_i - POČET PREDMETOV V i-TEJ PRIEHRADKE

VŠETKY STAVY SÚ ROVNAKO PRAVDEPODOBNÉ

POČET STAVOV $\binom{r+n-1}{r}$

$$P(K_i = k) = \frac{\binom{(r-k)+(n-1)-1}{r-k}}{\binom{r+n-1}{r}}$$

$r_n \to \infty, \quad \frac{r_n}{n} \to \lambda$

$$\frac{\binom{r-k+n-2}{n-2}}{\binom{r+n-1}{n-1}} = \frac{\frac{(r-k+n-2)\dots(n-1)}{(r-k)!}}{\frac{(r+n-1)\dots n}{r!}} =$$

$$= \frac{\overset{n-1}{\cancel{(n-1)!}}\, r!\, (r-k+n-2)!}{\cancel{(n-2)!}\,(r-k)!\,(r+n-1)!} = \frac{(n-1)(r-k+n-2)\dots(r-k+1)}{(r+n-1)\dots(r+1)} =$$

$$= \frac{n-1}{r+1} \cdot \frac{(r-k+n-2)\dots(r-k+1)}{(r+n-1)\dots(r+2)} \cdot \dots$$

$$\lim_{n\to\infty} P(K_i = k) = \overset{p}{\left(\frac{\lambda}{1+\lambda}\right)^k} \cdot \overset{1-p}{\frac{1}{1+\lambda}}$$

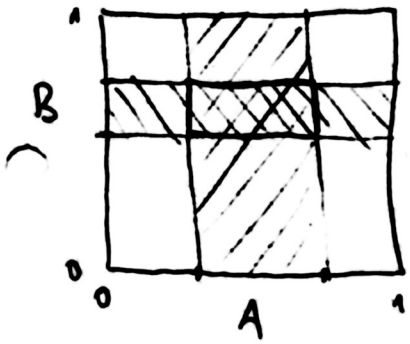
GEOMETRICKÉ ROZDELENIE

15.10.2007 M.
PETER ČERNO

MAI059 ~~ALGEBRA II~~ PRAVDEPODOBNOSŤ A ŠTATISTIKA

DEF.: JAV A A B SÚ NEZÁVISLÉ, KEĎ $P(A|B) = P(A)$

JAV A A B SÚ NEZÁVISLÉ, KEĎ $P(A \cap B) = P(A) \cdot P(B)$



DEF. $(\Omega, \mathcal{A}, P)$, $A_1, \dots, A_n$ SÚ JAV.
HOVORÍME, ŽE $A_1, \dots, A_n$ SÚ ZDRUŽENE NEZÁVISLÉ, KEĎ $\forall k \in \{2, \dots, n\}$:

(*) $P(A_{i_1} \cap \dots \cap A_{i_k}) = P(A_{i_1}) \cdot \dots \cdot P(A_{i_k})$
PRE $\forall \{i_1, \dots, i_k\} \in \binom{\{1, \dots, n\}}{k}$.

(TENTO PREDPOKLAD SA VEĽMI ŤAŽKO OVERUJE)

KEĎ (*) PLATÍ IBA PRE NIEKTORÉ PEVNÉ k, HOVORÍME, ŽE TIE JAV SÚ PO k-NEZÁVISLÉ.

VETA: KEĎ SÚ JAV NAVZÁJOM NEZÁVISLÉ (ZDRUŽENE), TOTOM SÚ PO k-NEZÁVISLÉ (PRE $\forall k \in \{2, \dots, n\}$).
NAOPAK TO NEPLATÍ.

PROTIPRÍKLAD: VEZMEME ŠTVORSTEN

1. STENA	ČERVENÁ	A
2.	MODRÁ	B
3.	BIELA	C
4.	ČERVENO-MODRO-BIELA	

$P(A) = \frac{2}{4}$, $P(B) = \frac{2}{4}$, $P(C) = \frac{2}{4}$, $P(A \cap B) = \frac{1}{4}$, $P(A \cap B \cap C) = \frac{1}{4}$!!

DEF. : NÁHODNÁ VELIČINA

$$X : (\Omega, \mathcal{A}, P) \mapsto (R_1, \mathcal{B})$$

TRIEDA MNOŽÍN V R_1

ALTERNATÍVNE ROZDELENIE

$$P(X=0) = p, \quad P(X=1) = 1-p$$

BINOMICKÉ ROZDELENIE S PARAMETRAMI n, p

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

GEOMETRICKÉ ROZDELENIE

$$P(X=k) = p(1-p)^k$$

NEGATÍVNE BINOMICKÉ ROZDELENIE

$$P(X=k) = \binom{k-1}{k-r} p^r (1-p)^{k-r}$$

ALEBO

$$P(X^*=k) = \binom{k+r-1}{r-1} p^r (1-p)^k$$

PREDPÍŠE SA NEJAKÝ VZOR Z, N (NAPR. ZZNZ) (ZDAR, NEZDAR)
AKÁ JE PRAVDEPODOBNOSŤ, ŽE SA DANÝ VZOR PRVÝKRÁT VYSKYTNE V k-TOM POKUSE PRI NEZÁVISLÝCH ALTERNATÍVNYCH POKUSOCH.

<u>POZRI SKRIPTÁ !</u>

ROVNOMERNÉ ROZDELENIE :

$$P(X=i) = \frac{1}{n} \qquad i = 1, \dots, n$$

16.10.2007 Tu
PETER ČERNO

MAI059 PRAVDEPODOBNOST A STATISTIKA

VETA O ÚPLNEJ PRAVDEPODOBNOSTI:

$$P(B) = \sum P(B|A_i)\,P(A_i) \qquad \Omega = \dot{\bigcup} A_i$$

BAYESOVA VETA:

$$P(A_j|B) = \frac{P(B|A_j)\,P(A_j)}{\sum P(B|A_k)\,P(A_k)}$$

PRÍKLAD: a ČIERNYCH GUĽOČIEK, b BIELYCH GUĽOČIEK TAHÁME BEZ VRACANIA. V 1. ŤAHU JE BIELA, AKÁ JE PRAVD. ŽE V 2. ŤAHU JE TIEŽ BIELA?

$$p = \frac{b-1}{a+b-1}$$

$$\Omega = \{\underset{A,\,B}{(B,B)},\ \underset{A}{(B,\bar{C})},\ \underset{B}{(\bar{C},B)},\ (\bar{C},\bar{C})\}$$

PRÍKLAD: 10 KLOBÚKOV $K_1, \dots, K_{10}$

$K_1, \dots, K_9 <$ 2 ČIERNE, 2 BIELE $\qquad K_{10} <$ 1 ČIERNA, 5 BIELYCH

a) NÁHODNE ZVOLÍM KLOBÚK A VYTIAHNEM BIELU. AKA JE PRAVD., ŽE SME ŤAHALI Z K_{10}

$$P(B|K_{1,\dots,9}) = \frac{1}{2}, \quad P(\bar{C}|K_{1,\dots,9}) = \frac{1}{2}$$

$$P(B|K_{10}) = \frac{5}{6}, \quad P(\bar{C}|K_{10}) = \frac{1}{6}$$

$$P(K_{10}|B) = \frac{P(B|K_{10}).P(K_{10})}{\sum P(B|K_i)\,P(K_i)} = \frac{\frac{5}{6}\cdot\frac{1}{10}}{9\cdot\frac{1}{2}\cdot\frac{1}{10} + \frac{5}{6}\cdot\frac{1}{10}} = \underline{\underline{\frac{5}{32}}}$$

OTEC MÁ SYNA, SYN HRÁ TENIS, OTEC PONÚKOL SYNOVI HODNOTNÝ DAR, KEĎ VYHRÁ 2 STRETNUTIA ZA SEBOU.

A: O - F - O

B: F - O - F

$P(O) = P(\text{VYHRÁ S OTCOM}) > P(F) = P(\text{VÝHRA S FEDERERom})$

	1	2	3	A	B
A_1	V	V	V	$p_O \cdot p_F \cdot p_O$	$p_F \cdot p_O \cdot p_F$
A_2	P	V	V	$(1-p_O)\, p_F \cdot p_O$	$(1-p_F) \cdot p_O \cdot p_F$
A_3	V	V	P	$p_O \cdot p_F \cdot (1-p_O)$	$p_F \cdot p_O \cdot (1-p_F)$
...					

$$P(DAR) = P(DAR \mid A_1)\, P(A_1) + \dots =$$
$$= P(A_1) + P(A_2) + P(A_3)$$

$$P_A(DAR) = p_O \cdot p_F\,(2-p_O)$$

$$P_B(DAR) = p_O\, p_F\,(2-p_F)$$

ZREJME $P_A(DAR) < P_B(DAR)$

16.10.2007 Štu
PETER ČERNO

MA1059 PRAVDEPODOBNOSŤ A ŠTATISTIKA

PRÍKLAD:

$$P(HIV \mid +) = \frac{P(+ \mid HIV) \cdot P(HIV)}{P(+ \mid HIV)\, P(HIV) + P(+ \mid Z) \cdot P(Z)}$$

$$P(HIV \mid ++) = \frac{P(+ \mid HIV) \cdot P(HIV \mid +)}{P(+ \mid HIV)\, P(HIV \mid +) + P(+ \mid Z) \cdot P(Z \mid +)}$$

A, B NEZÁVISLÉ $\Leftrightarrow$ $P(A \cap B) = P(A) \cdot P(B)$.
$P(A \mid B) = P(A)$

$A_1, \dots, A_n$ ZDRUŽENE NEZÁVISLÉ / k - NEZÁVISLÉ

PRÍKLAD: KLOBÚK, 4 LÍSTKY,
000, 110, 101, 011

A_i ... NA iTOM MIESTE TOHO LÍSTKU JE 1.

a) A_1, A_2, A_3 SÚ PO DVOJICIACH NEZÁVISLÉ? ÁNO
b) ZDRUŽENE NEZÁVISLÉ? NIE

PRÍKLAD: VÝHERNÝ AUTOMAT $\rightarrow$ $n-1$ MINCÍ
PÁN X $\rightarrow$ 1 MINCÍ

V KAŽDEJ HRE JEDEN Z NICH ZÍSKA A DRUHÝ STRATÍ MINCU.

$p = P(\text{VYHRÁ PÁN X})$
$q = P(\text{VYHRÁ AUTOMAT})$

$P(\text{PÁN X PREHRÁ ÚPLNE VŠETKO})$

OZNAČME $p_i = P(X$ VYHRÁ S VKLADOM $i)$

$p_0 = 0, \quad p_n = 1,$

$$p_i = a \cdot i + b$$

$$p_i = p \cdot p_{i+1} + q \cdot p_{i-1}$$

$$a \cdot i + b = p(ai + a + b) + q(ai - a + b)$$

$$ai = ai + pa - qa$$

$$p = q = \frac{1}{2} \quad ☹$$

$$p_i = \lambda^i \qquad \lambda^i = p\lambda^{i+1} + q\lambda^{i-1}$$

$$\lambda = p\lambda^2 + q$$

$$\lambda^2 - \frac{1}{p}\lambda + \frac{q}{p} = 0$$

$$\lambda_{1,2} = \frac{\frac{1}{p} \pm \sqrt{\frac{1}{p^2} - 4 \cdot \frac{q}{p}}}{2} = \frac{\frac{1}{p} \pm \sqrt{\frac{1}{p^2} - 4\,\frac{1-p}{p}}}{2} =$$

$$= \frac{\frac{1}{p} \pm \sqrt{\left(\frac{1}{p} - 2\right)^2}}{2}$$

PRE $p < \frac{1}{2}$ JE $\lambda_{1,2} = \dfrac{\frac{1}{p} \pm \left(\frac{1}{p} - 2\right)}{2}$ $\quad \boxed{\begin{aligned} \lambda_1 &= +1 \\ \lambda_2 &= \frac{1}{p} - 1 \end{aligned}}$

PRE $p = \frac{1}{2}$ MÁME RIEŠENIE $\boxed{p_i = \frac{i}{n}}$

PRE $p > \frac{1}{2}$ JE $\lambda_{1,2} = \dfrac{\frac{1}{p} \pm \left(2 - \frac{1}{p}\right)}{2}$ $\quad \begin{aligned} \lambda_1 &= +1 \\ \lambda_2 &= \frac{1}{p} - 1 \end{aligned}$

$$p_i = \alpha \cdot 1^{ai} + \beta \cdot \left(\frac{1}{p} - 1\right)^i \qquad \begin{aligned} \alpha + \beta &= 0 \\ \alpha + \beta\left(\frac{1}{p} - 1\right)^n &= 1 \end{aligned}$$

$$\alpha = \left[1 - \left(\frac{1}{p} - 1\right)^n\right]^{-1}$$

22.10.2007 Mo
PETER ČERNO

MAI059 PRAVDEPODOBNOSŤ A ŠTATISTIKA

BERNOULLIOVSKÁ PRAVD. NÁHODNÝ VÝBER S VRACANÍM

n NÁH. VELIČÍN $X_1, \dots, X_n \sim Alt(p)$

$X_i \in \begin{cases} 0 \\ 1 \end{cases}$ $\quad P(X_i = 1) = p \in \langle 0, 1 \rangle$

$P(X_i = 0) = 1 - p$

$$Y = \sum_{i=1}^{n} X_i \sim Bi(n, p)$$

$$P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad k = 0, 1, \dots, n$$

NÁHODNÝ VÝBER BEZ VRACANIA

(1.) $X_1, \dots, X_n$ NIE SÚ NEZÁVISLÉ

$P(X_i = 1) \neq P(X_j = 1)$

PRÍKLAD: MÁME N VÝROBKOV, M Z NICH JE POKAZENÝCH

n ROZSAH VÝBERU

$$|\Omega| = \binom{N}{n} \qquad P(Y = k) = \boxed{\frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}}}$$

$k \leq \min\{M, n\}$ DOLNÁ HRANICA

$k \geq \max\{0, n - (N - M)\}$ HORNÁ HRANICA

HYPERGEOMETRICKÉ ROZDELENIE $HPG(n, N, M)$

(2.)

A_1	p_1
$\vdots$	$\vdots$
A_m	p_m

$$\sum_{j=1}^{m} p_j = 1$$

ZAVEDIEME X_{ij} $\quad i=1,\dots,n$, $j=1,\dots,m$ $\quad$ OPAKOVANIE POKUSU

$X_{ij} < \begin{matrix} 1 \\ 0 \end{matrix}$ $\quad$ v i-TOM HODE NASTAL STAV A_j

$P(X_{ij}=1) = p_j \quad \forall j$

$\mathbb{X}_i = (X_{i1}, \dots, X_{im})^T$

$\mathbb{X}_i, \mathbb{X}_k$ SÚ NEZÁVISLÉ

$Y_j = \sum_{i=1}^{n} X_{ij} \rightarrow$ POČET A_j v n HODOCH

$\mathbb{Y} = (Y_1, \dots, Y_m)^T = \sum_{i=1}^{n} \mathbb{X}_i$

$$P(\mathbb{Y}=\mathbb{y}) = P(Y_1 = y_1, \dots, Y_m = y_m) =$$

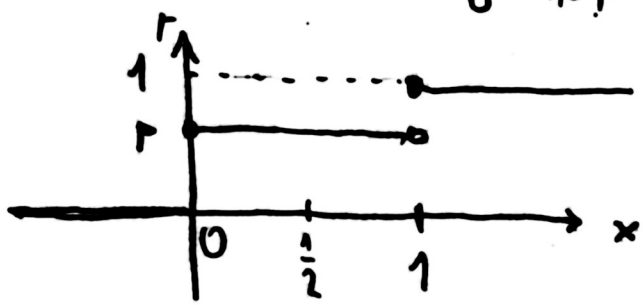
$$= \frac{n!}{y_1! \dots y_m!} p_1^{y_1} \dots p_m^{y_m}$$

$$\sum_{j=1}^{m} p_j = 1, \quad \sum_{j=1}^{m} y_j = n$$

$X_n \sim Bi(n, p_n) \quad p_n \in (0,1) \quad n p_n \rightarrow \lambda$

<u>DISTRIBUČNÁ FUNKCIA</u> $F_X(x) = P(X \leq x) = \sum_{s_i \leq x} p_i$

$X \sim Alt \quad X < \begin{matrix} 1 & p \\ 0 & 1-p \end{matrix}$



F JE NEKLESAJÚCA

S OBOROM HODNÔT $\langle 0,1 \rangle$

$\lim_{x \to +\infty} F(x) = 1, \quad \lim_{x \to -\infty} F(x) = 0.$

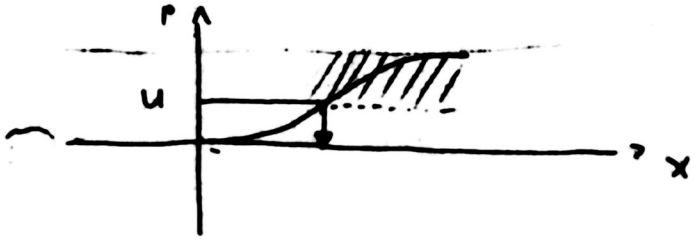
22.10.2007
PETER ČERNO

MA1059 PRAVD. A STATISTIKA

ZREJME $P(a < X \le b) = P(X \le b) - P(X \le a) =$
$= F(b) - F(a)$

FUNKCE KVANTILOVÁ

$F_{-1}(u) = \inf\{x \in \mathbb{R},\ F(x) > u\} \qquad u \in (0,1)$



$\exists f \quad f \ge 0, \quad \int_{\mathbb{R}} f(t)\,dt = 1$

$F(x) = \int_{-\infty}^{x} f(t)\,dt$

VYTVÁRAJÍCA FUNKCIA $A(x) = \sum_{i=0}^{\infty} p_i x^i$
$x \in \langle 0,1 \rangle$

$A(1) = 1, \quad A(0) = p_0$

$A'(x) = \sum_{i=1}^{\infty} i p_i x^{i-1} \qquad A'(1) = \sum_{i=1}^{\infty} i p_i = EX \qquad A'(0) = p_1$

$A''(x) = \sum_{i=2}^{\infty} i(i-1) p_i x^{i-2} \qquad A''(1) = \sum_{i=2}^{\infty} i(i-1) p_i, \quad A''(0) = 2p_2$

$A^{(n)}(0) = n!\, p_n \quad \Rightarrow \quad p_n = \frac{A^{(n)}(0)}{n!}$

$EX = \sum_{x_i \in \mathcal{X}} x_i P(X = x_i)$

$E(X - EX) < \begin{matrix} E(X-EX)^2 \\ E|X-EX| \end{matrix} \qquad \operatorname{var} X \qquad SD = \sqrt{\operatorname{var} X}$

TRANSF.

1) $X + c$ 2) aX 3) $X + Y$

23.10.2007 Tu
PETER ČERNO

EX: MAI059 PRAVDEPODOBNOSŤ A ŠTATISTIKA

NÁHODNÉ VELIČINY

$$X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

$$\forall B \in \mathcal{B}(\mathbb{R}) \Rightarrow X^{-1}(B) \in \mathcal{A}$$

$X(\omega)$

↑

NIEČO ABSTRAKTNÉ

ROZDELENIE X

$$\forall B \quad P(\omega : X(\omega) \in B) = P(X^{-1}(B))$$

STAČÍ URČIŤ PRAVDEPODOBNOSŤ

$$P(X^{-1}((-\infty, x\rangle)) = P(X \leq x) = F(x) \qquad \forall x \in \mathbb{R}$$

F JE DISTRIBUČNÁ FUNKCIA

F JE SPRAVA SPOJITÁ, NEKLESAJÚCA

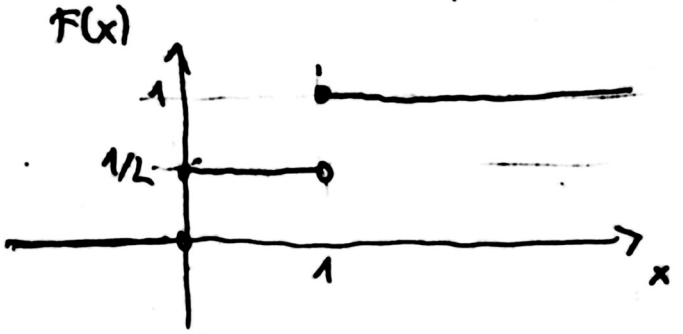
$$\lim_{x \to -\infty} F(x) = 0, \quad \lim_{x \to +\infty} F(x) = 1$$

NÁHODNÉ VELIČINY:

1) DISKRÉTNE → NADOBÚDA SPOČETNE VEĽA HODNÔT

$$P(X \leq x) = \sum_{x_i \leq x} P(X = x_i)$$

$\omega = (\text{POČASIE, TLAK, PRÍRODA}, \ldots) \to X(\omega)$



2) SPOJITÁ NÁHODNÁ VELIČINA

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

STREDNÁ HODNOTA

NIEČO AKO PRIEMERNÁ HODNOTA NÁH. VELIČINY

$$EX = \sum_{i=1}^{n} x_i p_i = \frac{\sum p_i x_i}{\sum p_i}$$

$$= \int_{-\infty}^{\infty} x f(x)\,dx = \frac{\int_{-\infty}^{\infty} x f(x)\,dx}{\int_{-\infty}^{\infty} f(x)\,dx}$$

PRÍKLAD: ŠTUDENT → TEST S OTÁZKAMI (5 OTÁZOK)
VYBERÁ Z 3 MOŽNOSTÍ, IBA 1 JE SPRÁVNE

X = NÁH. VELIČINA, POČET SPRÁVNE ZODPOVEDANÝCH OTÁZOK

1) POPÍSAŤ ROZDELENIE N. V. X.

1 OTÁZKA $Y = \begin{cases} 0 & P(Y=0) = \frac{2}{3} \\ 1 & P(Y=1) = \frac{1}{3} \end{cases}$

$$X \in \{\underset{p_0}{0}, \underset{p_1}{1}, \dots, \underset{p_5}{5}\} \qquad P(X=i) =: p_i$$

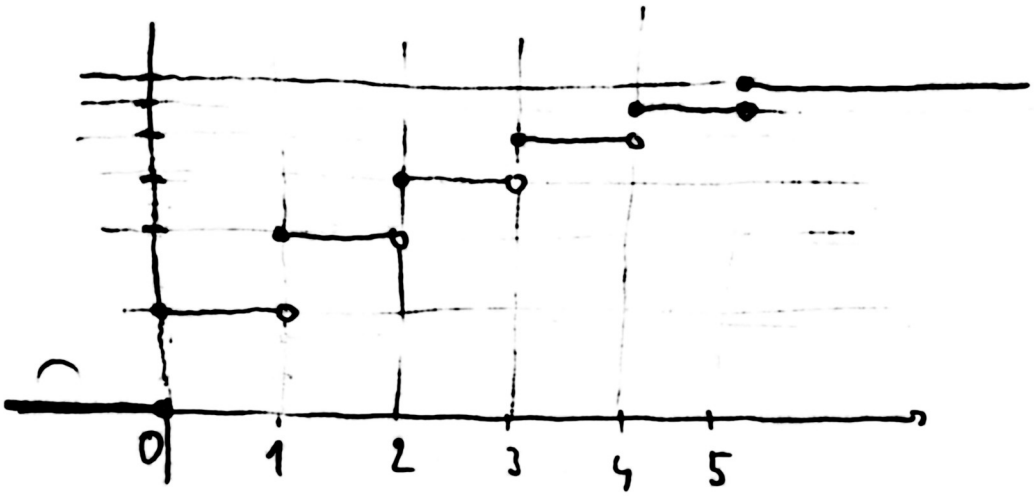
$$p_0 = \left(\frac{2}{3}\right)^5, \quad p_1 = \binom{5}{1}\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^4, \dots, p_5 = \binom{5}{5}\left(\frac{1}{3}\right)^5\left(\frac{2}{3}\right)^0$$

TZV. BINOMICKÉ ROZDELENIE.

23.10.2007 Tu
PETER ČERNO

EX: MAI059 PRAVD. A STATISTIKA

DISTRIBUČNÁ FUNKCIA



ÚSPECH U SKÚŠKY = ASPOŇ 4 OTÁZKY SPRÁVNE

$$EX = \frac{5}{3}$$

$$\sum_{k=0}^{n} k \cdot P(X=k) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k} = p \cdot n$$

PRÍKLAD:

STROJ NA GENEROVANIE NÁHODNÝCH ČÍSIEL.
GENERUJE ČÍSLICE 0 .. 9 (KAŽDÚ S ROVNAKOU PRAVD.)
ČÍSIEL GENERUJEME n.
AKÝM ROZDELENÍM SA RIADI POČET PÁRNYCH ČÍSIEL?

$$X \in \{0, 1, 2, \dots, n\}$$

$$P(X=k) = \binom{n}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{n-k} = \binom{n}{k} \frac{1}{2^n}$$

BINOMICKÉ ROZDELENIE

$$EX = \frac{n}{2}$$

KOĽKO ČÍSIEL MUSÍM GENEROVAŤ, ABY S PRAVDEPODOBNOSŤOU ASPOŇ 0,975 DOSTAL 1 ALEBO VIAC SUDÝCH ČÍSIEL?

$$P(X \geq 1) = 1 - P(X=0) = 1 - \frac{1}{2^n} \geq 0{,}975$$

$$n \geq \left\lceil \log_2 \frac{1}{1-0{,}975} \right\rceil = \left\lceil -\log_2(0{,}025) \right\rceil =$$

$$= \lceil \log_2(40) \rceil = 6$$

PODOBNE

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - \left(\frac{1}{2^n} + \binom{n}{1}\frac{1}{2^n}\right) =$$

$$= 1 - \frac{n+1}{2^n} \geq 0{,}975$$

$$\frac{n+1}{2^n} \leq 0{,}025 \qquad n \geq 9$$

PRÍKLAD: 2 HRÁČI BASKETBALU HÁDŽU NA KOŠ, STRIEDAJÚ SA, HÁDŽU DO TEJ DOBY, DOKEDY JEDEN Z NICH NEDOSIAHNE KOŠ.

A → PRST ÚSPEŠNÉHO HODU = 0,4 p
B → —"— = 0,6 q

a) ROZDELENIE POČTU HODOV, KT. DOSIAHNE KAŽDÝ Z NICH

X = POČET HODOV HRÁČA A $\{1, 2, \dots\}$
Y = POČET HODOV HRÁČA B $\{0, 1, \dots\}$

$$P(X=1) = 0{,}4 + (1-0{,}4) \cdot 0{,}6$$

$$P(X=n) = [(1-0{,}4)\cdot(1-0{,}6)]^{n-1} \cdot 0{,}4 + [(1-0{,}4)\cdot(1-0{,}6)]^{n-1} \cdot (1-0{,}4) \cdot 0{,}6 = (0{,}4 \cdot 0{,}6)^{n-1}(0{,}4 + 0{,}6 \cdot 0{,}6)$$

23.10.2007 Tu
PETER ČERNO

EX: MAI059 PRAVD. A STATISTIKA

$$P(Y=n) = [(1-0{,}4)\cdot(1-0{,}6)]^{n-1}\cdot(1-0{,}4)\cdot 0{,}6 +$$
$$+ [(1-0{,}4)\cdot(1-0{,}6)]^{n-1}\cdot(1-0{,}4)\cdot(1-0{,}6)\cdot 0{,}4 =$$
$$= [0{,}6\cdot 0{,}4]^{n-1}(0{,}6\cdot 0{,}6 + 0{,}6\cdot 0{,}4\cdot 0{,}4)$$

$$P(X=k) = (1-r)^{k-1} r \quad \text{PRE} \quad r = p+q-pq$$

GEOMETRICKÉ ROZDELENIE

$Z = X+Y$?

$P(Z=n)$?

$$\sum x^i = \frac{1}{1-x}$$
$$\sum i x^{i-1} = \frac{1}{(1-x)^2}$$
$$\sum i(1-x)^{i-1} x = \frac{1}{x}$$

$$P(Z=1) = p$$
$$P(Z=2) = (1-p)q$$
$$P(Z=3) = (1-p)(1-q)p$$
$$\dots$$
$$P(Z=2k+1) = [(1-p)(1-q)]^k p$$
$$P(Z=2k+2) = [(1-p)(1-q)]^k (1-p) q$$

$$EZ = E(X+Y) = EX + EY \qquad EY = \frac{1-p}{r}$$

$$\text{"}E(\text{Geom}(p)) = \frac{1-p}{p}\text{"} \qquad \boxed{\sum_{k=0}^{\infty} k(1-p)^k p = \frac{1-p}{p}}$$

$$EX = \sum_{k=1}^{\infty} k\cdot P(X=k) = \sum_{k=1}^{\infty} k\cdot(1-r)^{k-1} r =$$
$$= \underbrace{\sum_{k=1}^{\infty}(k-1)(1-r)^{k-1}\cdot r}_{\frac{1-r}{r}} + \underbrace{\sum_{k=1}^{\infty}(1-r)^{k-1}\cdot r}_{r\cdot\frac{1}{1-(1-r)}} = \frac{1}{r}$$

29.10.2007 Mb
PETER ČERNO

MA1059 PRAVDEPODOBNOST A ŠTATISTIKA

$$\boxed{E}X = \boxed{\sum \boxed{x_i}\boxed{p_i}} \qquad EX^2 = \boxed{\sum \boxed{x_i^2}\boxed{p_i}}$$

$$= \boxed{\int_{-\infty}^{\infty} \boxed{x}\boxed{f(x)\,dx}} \qquad = \boxed{\int_{-\infty}^{\infty} \boxed{x^2}\boxed{f(x)\,dx}}$$

$$Ea = \sum a p_i = a\left(\sum p_i\right) = a$$

$$E(aX) = a\,EX$$

$$E(X+Y) = \sum_i \sum_j (x_i + y_j)\, P(X=i, Y=j) =$$

$$= \sum_i x_i \sum_j P(X=i, Y=j) + \sum_j y_j \sum_i P(X=i, Y=j)$$

$$P\left(\bigcup_j \{X=i, Y=j\}\right) = P(X=i)$$

$$= \sum_i x_i P(X=i) + \sum_j y_j P(Y=j) = EX + EY.$$

PRÍKLAD:

$A = \{OS_1$ ... SPADNE POČAS 6-TICH POZÁRUČNÝCH MESIACOCH$\}$
$B = \{OS_2$ ... -||- $\}$
$C = \{OS_3$ ... -||- $\}$

$P(A) = 0{,}4, \quad P(B) = 0{,}1, \quad P(C) = 0{,}2$

100 POBOČIEK, V KAŽDEJ SÚ 3 POČÍTAČE, 1 OS/PC
JEDNOTLIVE PC SÚ NA SEBE NEZÁVISLÉ

? OČAKÁVANÝ POČET VÝJAZDOV ZDARMA

$A_1, \dots, A_{100} \sim Alt(0,4)$

$B_1, \dots, B_{100} \sim Alt(0,1)$

$C_1, \dots, C_{100} \sim Alt(0,2)$

$$EA_i = \sum_{i=0}^{1} x_i p_i = 0 \times p_0 + 1 \times p_1 = 0,4$$

$$\boxed{S_A = \sum_{i=1}^{100} A_i} \sim Bi(100, 0.4)$$

$$ES_A = E\left(\sum_{i=1}^{100} A_i\right) = \sum_{i=1}^{100} EA_i = 100 \cdot 0,4 = 40$$

$$ES_A = \sum_{j=0}^{100} \binom{100}{j} 0,4^j \cdot (1-0,4)^j \cdot j = 0,4 \cdot 100 = 40.$$

$ES_B = 10$

$ES_C = 20$

PRÍKLAD:

X	...	0	1	2	...
p		p_0	p_1	p_2	...

$$A(t) = \sum_{i=0}^{\infty} p_i t^i$$

$X \sim Alt(p)$ $\quad X < \begin{smallmatrix} 1 \\ 0 \end{smallmatrix}$ $\quad A(t) = p_0 \cdot t^0 + p_1 t^1 = q + pt$

$X \sim Bi(n,p)$ $\quad A(t) = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} \cdot t^i = (q+pt)^n$

$X \sim Ge(p)$

$P(X=j) = p \cdot q^j$ $\quad A(t) = \sum_{j=0}^{\infty} p q^j t^j =$

$$= p \sum_{j=0}^{\infty} (qt)^j = p \cdot \frac{1}{1-qt}$$

29.10.2007 Mo
PETER ČERNO

MA105g PRAVDĚPODOBNOST A STATISTIKA

ZREJME $EX = A'(1)$.

$A_{Alt}(t) = q + pt$ $\qquad$ $A'_{Alt}(1) = p$

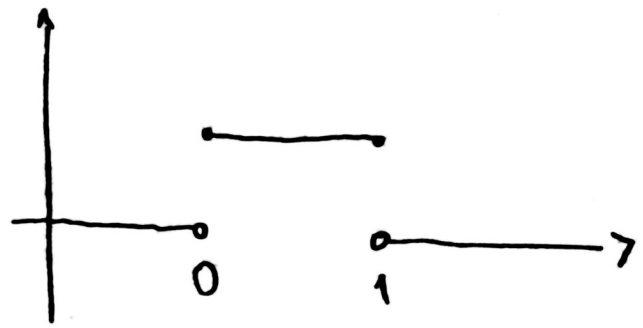
$A_{Bin}(t) = (q+pt)^n$ $\qquad$ $A'_{Bin}(t) = n(q+pt)^{n-1}p \Rightarrow A'_{Bin}(1) = np$

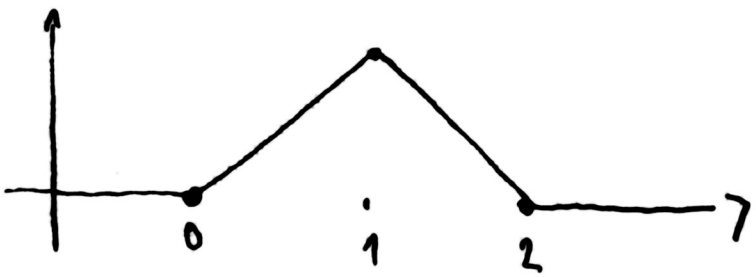
$A_{Geom}(t) = \frac{p}{1-qt}$ $\qquad$ $A'_{Geom}(t) = \frac{p\cdot(-1)\cdot(-q)}{(1-qt)^2} = \frac{pq}{(1-qt)^2}$

$\Rightarrow A'_{Geom}(1) = \frac{q}{p}$

X, Y NEZÁVISLÉ, POTOM

$$A_{X+Y}(t) = A_X(t)\cdot A_Y(t).$$





$\operatorname{var} X = E\,(X-EX)^2$ VARIANCE

$E\,|X-EX|$ STREDNÍ ABSOLUTNÍ CHYBA

$$\operatorname{var} X = E\,(X-EX)^2 = E(X^2 - 2X(EX) + (EX)^2) =$$
$$= E(X^2) - 2EX\,.\,(EX) + E(EX)^2$$
$$= E(X^2) - (EX)^2 \geq 0$$ CAUCHY - SCHWARZ

$$\operatorname{var}(a+X) = E\,(X+a - E(X+a))^2 = \operatorname{var} X.$$

$$\operatorname{var}(aX) = a^2 \operatorname{var}(X)$$

SD, std, st.dev STANDARD DEVIATION

$$\underline{\underline{\operatorname{st}(X) \overset{\text{def}}{=} \sqrt{\operatorname{var}(X)}}}$$

$$\operatorname{var}(X+Y) = \operatorname{var} X + \operatorname{var} Y + 2\operatorname{cov}(X,Y)$$
$$\operatorname{cov}(X,Y) = E(X-EX)(Y-E(Y))$$

X, Y NEZÁVISLÉ $\Rightarrow$ $\operatorname{cov}(X,Y) = 0$

30.10.2007 Tu
PETER ČERNO

EX: MA1059 PRAVD. A STATISTIKA

$X = 1000 \cdot i$ KČ s PRAVD. $\frac{1}{2^i}$ $i = 1,2,3$
$= 0$ KČ s PRAVD. $\frac{1}{8}$

$$EX = \sum_{k=0}^{\infty} x_k \cdot p_k = \frac{1}{2} 1000 + \frac{1}{4} 2000 + \frac{1}{8} 3000 = \underline{\underline{1375}}$$

PRÍKLAD: POČET SPAMOV ZA DEŇ SA RIADI POISSONOVÝM ROZDELENÍM $Poiss(\lambda)$.
STREDNÝ POČET SPAMOV ZA DEŇ JE 25.

a) ČOMU JE ROVNÉ λ ?

b) AKÁ JE PRAVDEPODOBNOSŤ, ŽE DNES STRÁVIM MAZANÍM SPAMOV NAJVIAC 1,5 min., KEĎ ZMAZANIE SPAMU TRVÁ 3s.

$X \sim Poiss(\lambda)$ $\quad X \in \{0, 1, 2, \dots\}$

$$P(X=k) = \frac{e^{-\lambda} \lambda^k}{k!} \qquad EX = \lambda$$

a) $\lambda = 25$

b) $1{,}5\,min = 90\,s \rightsquigarrow 30$ SPAMOV

$$P(X \leq 30) = \sum_{k=0}^{30} \frac{e^{-\lambda} \lambda^k}{k!} \sim 0{,}8633$$

$$\boxed{\psi(t) = E e^{tX} = \sum_{k=0}^{\infty} e^{t x_k} p_k}$$

MOMENTOVÁ VYTVÁRAJÚCA FUNKCIA

$\psi'(0) = EX$

$\psi''(0) = EX^2$ $\qquad$ $\operatorname{var} X = E(X - EX)^2 = EX^2 - (EX)^2$

$$EX^2 = \sum_{k=0}^{\infty} k^2 P(X=k) = \sum_{k=0}^{\infty} k^2 \cdot \frac{e^{-\lambda}\lambda^k}{k!} =$$

$$= \frac{e^{-\lambda}\cdot\lambda}{1!} + \lambda \sum_{k=2}^{\infty} k \cdot \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} =$$

$$= \frac{e^{-\lambda}\cdot\lambda}{1!} + \lambda\left(\sum_{k=1}^{\infty} k \frac{e^{-\lambda}\lambda^k}{k!}\right) + \lambda\left(\sum_{k=2}^{\infty} \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!}\right) =$$

$$= e^{-\lambda}\cdot\lambda + \lambda\cdot\lambda + \lambda e^{-\lambda}\left[\left(\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\right) - 1\right] =$$

$$= e^{-\lambda}\cdot\lambda + \lambda^2 + \lambda e^{-\lambda}(e^{\lambda} - 1) =$$

$$= \underline{\underline{\lambda^2 + \lambda}}$$

$$\psi(t) = E e^{tX} = \sum_{k=0}^{\infty} e^{tk} \cdot P(X=k) = \sum_{k=0}^{\infty} e^{tk} \cdot \frac{e^{-\lambda}\lambda^k}{k!}$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{k!} = \underline{\underline{e^{-\lambda} e^{\lambda e^t}}}$$

$$\psi'(t) = e^{-\lambda} \cdot \lambda e^t \cdot e^{\lambda e^t}$$

$$\psi''(t) = \lambda e^{-\lambda}\left[e^t \cdot e^{\lambda e^t} + e^t \cdot \lambda e^t e^{\lambda e^t}\right] =$$

$$= \lambda e^{-\lambda} \cdot e^t \cdot e^{\lambda e^t} \cdot (1 + \lambda e^t)$$

$$\psi''(0) = \lambda \cdot 1 \cdot 1 \cdot 1 \cdot (1 + \lambda \cdot 1) = \underline{\underline{\lambda + \lambda^2}}$$

30.10.2007 Tu
PETER ČERNO

EX: MAJOR PRAVDEPODOBNOSŤ A ŠTATISTIKA

$$\boxed{E\,f(x) = \sum_{k=0}^{\infty} f(x_k)\, P(X = x_k)}$$

PRÍKLAD: NA ÚZEMÍ LONDÝNA DOPADLO 537 RAKIET V1 A V2. ROZDELÍME LONDÝN NA $24^2 = 576$ ROVNAKO VEĽKÝCH OBLASTÍ.
AKÁ JE PRIBLIŽNE PRAVDEPODOBNOSŤ, ŽE NA KONKRÉTNY ~1 ŠTVOREC DOPADNE k RAKIET, ZA PREDPOKLADU, ŽE RAKETY NIE SÚ ŠPECIÁLNE ZAMEROVANÉ NA JEDNOTLIVÉ OBLASTI.

$$P(\text{ZÁSAH}) = \frac{1}{576}$$

$$P(X=k) = \binom{537}{k} \cdot \left(\frac{1}{576}\right)^k \cdot \left(1 - \frac{1}{576}\right)^{537-k}$$

BINOMICKÉ ROZDELENIE $\text{Binom}\left(537, \frac{1}{576}\right)$

$n = 537, \quad p = \frac{1}{576}$

SÚVISLOSŤ MEDZI BINOMICKÝM A POISSONOVÝM ROZDELENÍM

$$n p_n \to \lambda$$

$$P(X=k) \to \frac{e^{-\lambda}\lambda^k}{k!}$$

n_i ... POČET RAKIET, KT. SPADLI DO i-TEJ OBLASTI

$n_i \in \{0, \dots, 537\}$

$$m_0 = \frac{\# i : n_i = 0}{576} \rightarrow \text{ODHAD PRE } P(X=0)$$

...

$$m_r = \frac{\# i : n_i = r}{576} \rightarrow \text{ODHAD PRE } P(X=r)$$

PRE ZAUJÍMAVOSŤ:

# RAK.	0	1	2	3	4	≥ 5
# OBL.	229	211	93	35	7	1
TEÓRIA	226,7	211,4	98,5	30,6	7,1	...

$576 \cdot P(X=k)$

ZHRNUTIE: DISKRÉTNE ROZDELENIA:

a) ALTERNATÍVNE ROZDELENIE

MODELUJE HÁDZANIE KORUNOU $(p, 1-p)$

b) BINOMICKÉ ROZDELENIE

OPAKOVANÝ HOD KORUNOU

c) POISSONOVÉ ROZDELENIE

POČTY NIEČOHO $EX = \lambda$, $P(X=k) = \frac{e^{-\lambda} \lambda^k}{k!}$

d) GEOMETRICKÉ ROZDELENIE

POČET NEÚSPECHOV PRED 1. ÚSPECHOM

30.10.2007 T4
PETER ČERNO

EX: MAIOJ PRAVDEPODOBNOSŤ A ŠTATISTIKA

e) NEGATÍVNE BINOMICKÉ ROZDELENIE
POČET NEÚSPECHOV PRED m-TÝM ÚSPECHOM

f) ~~DISKRÉTNE ROZDELENIE~~
~~ROVNOMERNÉ ROZDELENIE~~

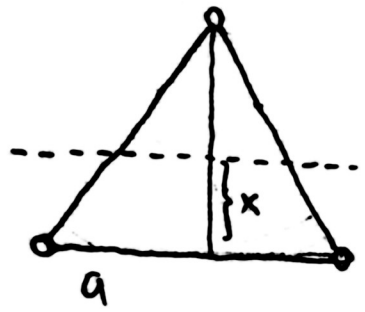
HYPERGEOMETRICKÉ ROZDELENIE
VÝBER BEZ VRACIANIA Z KLOBÚKU
M BIELYCH, N ČIERNYCH GUĽ

PRÍKLAD: LES MÁ TVAR ROVNOSTRANNÉHO Δ.
V LESE SA STRATILA ČERVENÁ KARKULKA.
PRAVDEPODOBNOSŤ, ŽE Č.K. JE V NEJAKEJ ČASTI LESA JE PRIAMO ÚMERNÁ VEĽKOSTI TEJTO ČASTI, NEZÁVISÍ NA UMIESTNENÍ.
VĽK STOJÍ NA PEVNE ZVOLENOM KRAJI LESA.
AKÉ JE ROZDELENIE VZDIALENOSTI MEDZI ČERVENOU KARKULKOU A STRANOU LESA S VĽKOM

$$S_\Delta = \frac{a^2\sqrt{3}}{4} \qquad v = a\frac{\sqrt{3}}{2} \qquad S = \frac{v^2\sqrt{3}}{3}$$

ROZDELENIE POPÍSEME DISTRIBUČNOU FUNKCIOU

$$F(x) = P(X \leq x)$$



$$F(x) = \frac{S_\Delta - \frac{(v-x)^2\sqrt{3}}{3}}{S_\Delta} =$$

$$= 1 - \frac{\left(\frac{\sqrt{3}}{2}a - x\right)^2\sqrt{3}}{\sqrt{3}\,\frac{3}{4}a^2} = 1 - \frac{\frac{3}{4}a^2 - \sqrt{3}ax + x^2}{\frac{3}{4}a^2}$$

$$F(x) = 1 - \left(1 - \frac{4}{3}\sqrt{3}\,\frac{x}{a} + \frac{4}{3}\,\frac{x^2}{a^2}\right) =$$

$$= \underline{\underline{4\left(\frac{x}{\sqrt{3}\,a} - \frac{x^2}{3a^2}\right)}} \qquad x \in \langle 0;\ \tfrac{\sqrt{3}}{2}a \rangle$$

$$F(x) = \int_{-\infty}^{x} f(t)\,dt \quad \Rightarrow \quad f(x) = F'(x)$$

$$F'(x) = \underline{\underline{4\left(\frac{1}{\sqrt{3}a} - \frac{2x}{3a^2}\right)}} = f(x)$$

$$EX = \int_{0}^{\frac{\sqrt{3}}{2}a} x\,f(x)\,dx =$$

$$= \int_{0}^{\frac{\sqrt{3}}{2}a} 4\left(\frac{x}{\sqrt{3}a} - \frac{2x^2}{3a^2}\right)dx =$$

$$= 4\left[\frac{x^2}{2\sqrt{3}a} - \frac{2x^3}{9a^2}\right]_0^{\frac{\sqrt{3}}{2}a} = 4\underbrace{\left[\frac{\frac{3}{4}a^2}{2\sqrt{3}a} - \frac{2\cdot\frac{\sqrt{3}\,3}{8}a^3}{9a^2}\right]}_{\frac{\sqrt{3}}{8}a - \frac{\sqrt{3}}{3}\cdot\frac{1}{4}a = \frac{\sqrt{3}}{24}a} =$$

$$= \underline{\underline{a\,\frac{\sqrt{3}}{6}}}$$

AKÉ JE ROZDELENIE VZDIALENOSTI KARKULKY OD NAJBLIŽŠEJ STRANY LESA ?

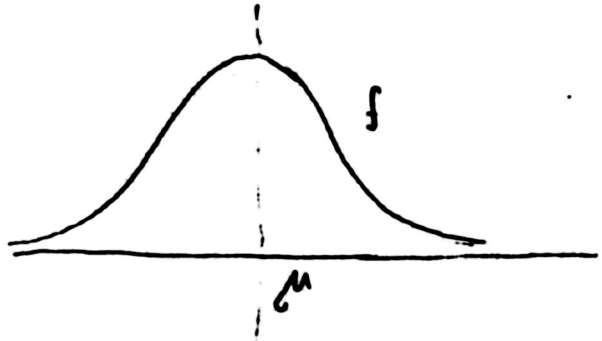
19.11.2007
PETER ČERNO

MA1059 PRAVDEPODOBNOSŤ A ŠTATISTIKA

GAUSSOVO - NORMÁLNE ROZDELENIE

$$f(x, \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$x \in \mathbb{R}$,
$\mu \in \mathbb{R}$,
$\sigma^2 > 0$



$X \sim N(\mu, \sigma^2)$

$EX = \mu$

$\operatorname{var} X = \sigma^2$

ŠTANDARDIZOVANÉ N. R. $N(0,1)$ $\quad f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

$X \sim N(0,1) \Rightarrow Y = \mu + \sigma X \sim N(\mu, \sigma^2)$

VETA: KEĎ $X \sim (\mu, \sigma^2)$ A $Y \sim N(\nu, \tau^2)$ A SÚ NEZÁVISLÉ, POTOM

$$aX + bY \sim N(a \cdot \mu + b \cdot \nu,\ a^2\sigma^2 + b^2\tau^2)$$

DÔKAZ. POMOCOU CHARAKT. FCIE. $\varphi(t) = Ee^{itX}$

VETA PLATÍ AJ V PRÍPADE, KDE X A Y SÚ ZÁVISLÉ, ALE

$$aX + bY \sim N(a\mu + b\nu,\ a^2\sigma^2 + b^2\tau^2 + a \cdot b \cdot \operatorname{cov}(X,Y))$$

PRÍKLAD: $X \sim N(1,1)$, $Y \sim N(2,4)$, POTOM

$X - Y \sim N(-1,5)$.

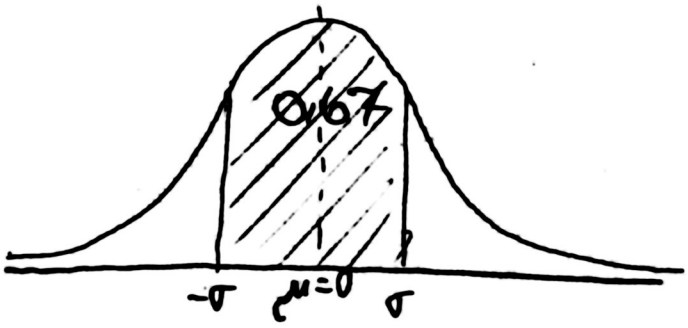
VETA: PRE NORMÁLNE ROZDELENIE PLATÍ

$X \sim N(0, \sigma^2)$

$P(-\sigma < X < \sigma) \approx 0{,}67$

$P(-2\sigma < X < 2\sigma) \approx 0{,}95$

$P(-3\sigma < X < 3\sigma) \approx 0{,}997$



POZNÁMKA: EXISTUJÚ TABUĽKY DISTRIBUČNEJ A KVANTILOVEJ FUNKCIE

<u>CENTRÁLNA LIMITNÁ VETA</u>: (MOIVRE - LAPLACEOVA)

NECH $X_1, X_2, \dots$ NEZÁVISLÉ, ALTERNATÍVNE ROZDELENÉ N. V.. POTOM

$$P\left(\frac{\sum_{i=1}^{n} X_i - np}{\sqrt{np(1-p)}} \leq x\right) \approx \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$X_i < \begin{matrix} 1 & p \\ 0 & 1-p \end{matrix}$$

$\sum_{i=1}^{n} X_i \sim Bi(n,p)$

$E \sum_{i=1}^{n} X_i = np$

$var\left(\sum_{i=1}^{n} X_i\right) = np(1-p)$

19.11.2007 Mo
PETER ČERNO

MA1059 PRAVD. A STAT.

$$P\left(\frac{Bi(n,p) - E\,Bi(n,p)}{\sqrt{var\,Bi(n,p)}} \le x\right) \approx \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,dt$$

$$P\left(\frac{\sum_{i=1}^{n} X_i - E\sum_{i=1}^{n} X_i}{\sqrt{var\,\sum_{i=1}^{n} X_i}} \le x\right) \approx \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,dt$$

VETA: NECH X_i SÚ NEZÁVISLÉ, ROVNAKO ROZDELENÉ, $EX_i^2 < \infty$, POTOM PLATÍ

PODSTATNÉ: KEĎ SÚČAT NÁHODNÉ VELIČINY (ALEBO PRIEMERUJEM) $\Rightarrow$ VEĽMI RÝCHLO MÔŽEM NAMIESTO PREJNÉHO ROZDELENIA POUŽIŤ NORMÁLNOU APROXIMÁCIU.

X .. $0, 1, 2, \dots$ $\quad p_i$
Y .. $0, 1, 2, \dots$ $\quad q_i$
$X+Y$ .. $0, 1, 2, \dots$ $\quad \sum_{j=0}^{i} p_j q_{i-j}$

MOŽNO APROXIMOVAŤ

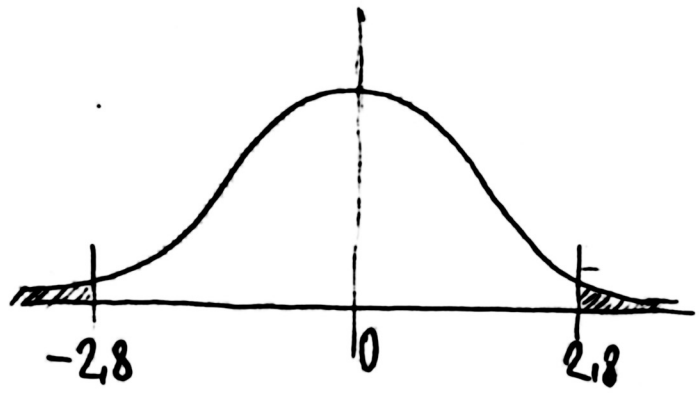
SKÚSENOSŤ: TÁTO APROXIMÁCIA JE DOBRÁ AJ PRE VEĽMI MALÉ n.

PRE $X \sim Alt$ $\quad p \in (0{,}1;\ 0{,}9)$
$n \approx 50$

PRE $X \sim R(0,1)$ $\quad n \approx 12$.

ALGORITMUS PRE GENEROVANIE ČÍSIEL Z $N(0,1)$

```
S = 0
FOR i = 1 TO 12
    S = S + RND
END
RETURN S - 6
```



$$ES = E\sum_{i=1}^{12} R_i = 6$$

$$\operatorname{var} R_i = \frac{1}{12} \Rightarrow \operatorname{var} S = 1.$$

HLAVNÝ VÝZNAM: JE V APROXIMÁCII ROZDELENIA PRIEMERU

$$\overline{X}_n \to \mu = EX_1$$

$X_1, \dots, X_n$ ... NÁHODNÉ POKUSY

$$P(|\overline{X}_n - \mu| \le \varepsilon) \ge 1 - \alpha \doteq 0{,}99$$

KOĽKO MÁM UROBIŤ POKUSOV? POUŽIJEME $\mu = 0$

$$P(|\overline{X}_n| \le \varepsilon) = P(-\varepsilon \le \overline{X}_n \le \varepsilon) =$$

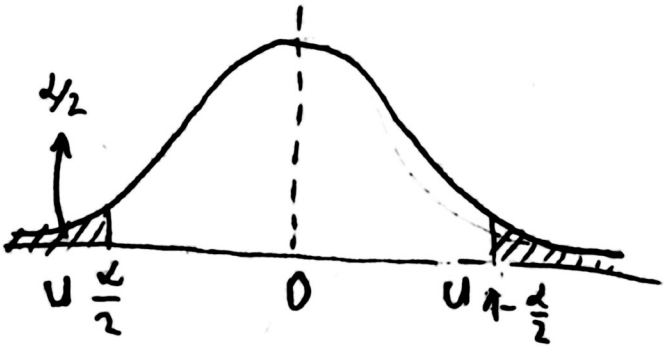
$$= P(-n\varepsilon \le \sum X_i \le n\varepsilon)$$

$$P\Bigg(\underbrace{\frac{-n\varepsilon - \overset{=0}{E\sum X_i}}{\sqrt{\operatorname{var}\sum X_i}}}_{\frac{-n\varepsilon}{\sqrt{n\sigma^2}} = -\sqrt{n}\frac{\varepsilon}{\sigma}} \le \underbrace{\frac{\sum X_i - \overset{=0}{E\sum X_i}}{\sqrt{\operatorname{var}\sum X_i}}}_{\sim N(0,1)} \le \underbrace{\frac{n\varepsilon - \overset{=0}{E\sum X_i}}{\sqrt{\operatorname{var}\sum X_i}}}_{\sqrt{n}\frac{\varepsilon}{\sigma}}\Bigg)$$

19.11.2007/6
PETER ČERNO

MA1059 PRAVD. A STAT.

$$= \Phi\left(\sqrt{n}\,\frac{\varepsilon}{\sigma}\right) - \Phi\left(-\sqrt{n}\,\frac{\varepsilon}{\sigma}\right).$$



$$u_t = \Phi^{-1}(t)$$

$$\sqrt{n}\,\frac{\varepsilon}{\sigma} = u_{1-\frac{\alpha}{2}} \quad \Rightarrow \quad n = \frac{u^2_{1-\frac{\alpha}{2}}\,\sigma^2}{\varepsilon^2}$$

ZÁVER: SIMULÁCIA NIE JE VHODNÁ METÓDA PRE ZÍSKANIE PRESNÝCH VÝSLEDKOV.

PRÍKLAD: O PREVOZNÍKOVI

NOSNOSŤ 5000 kg

KOĽKO MÔŽEM PREDAŤ LÍSTKOV, ABY SOM PRÁV NEPREVÁŽIL O VIAC NEŽ 100 kg SO SPOĽAHLIVOSŤOU $1-\alpha$.

X_i ~ VÁHA i-TEHO ČLOVEKA

$\sum X_i$ ZAŤAŽENIE PRÁVU

$$P\left(\sum^{n} X_i > \underset{5100}{5000}\right) \leq \alpha$$

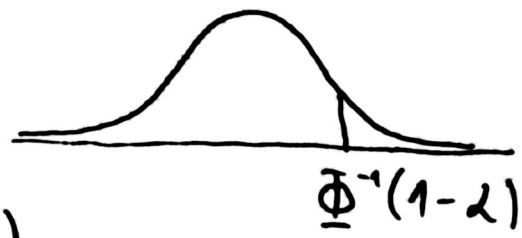
NEPOZNÁME POZNÁME

$X \sim N(70, 20^2)$

$70 \pm 3 \cdot 20 \equiv [10, 130]$ .. 99,7%

$\Rightarrow \sum X_i \sim N(70n, 400n)$

$$p\left(\underbrace{\frac{\sum X_i - 70n}{20\sqrt{n}}}_{N(0,1)} > \frac{5000 - 70n}{20\sqrt{n}}\right) = \alpha$$



$$\frac{5000 - 70n}{20\sqrt{n}} = \Phi^{-1}(1-\alpha)$$

ČO JE KVADRATICKÁ ROVNICA PRE n.

VÍDE PRIBLIŽNE $n = 64$.

ÚLOHA: (NA ZAMYSLENIE)

AKO STANOVIŤ K, ABY SME PRI #ZDAROV $< K$ ALEBO PRI #ZDAROV $> K$ HO MOHLI (S ČISTÝM SVEDOMÍM) PREHLÁSIŤ ZA PODVODNÍKA.

05.11.2007 Mo
PETER ČERNO

MA1059 PRAVDEPODOBNOSŤ A ŠTATISTIKA

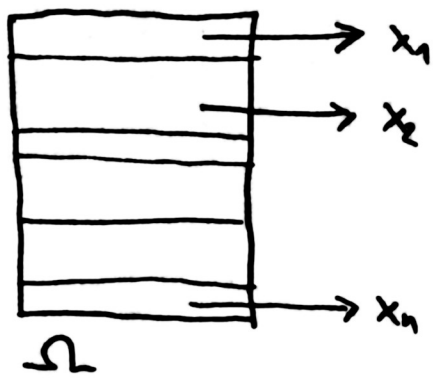
$$X : (\Omega, \mathcal{A}, P) \to (R_1, \dots)$$

$$EX = \sum_{\omega} X(\omega) \cdot P(\omega)$$

$$X : \omega \to X(\omega) = x_i \in R_1$$

KONEČNE / SPOČETNE VEĽA

$$EX = \sum_i x_i \cdot p_i$$



$$P(X = x_i) = \sum_{\substack{\omega \\ X(\omega) = x_i}} P(\omega)$$

$$EX = \sum_i \sum_{\substack{\omega \\ X(\omega) = x_i}} \overbrace{X(\omega)}^{x_i} \cdot P(\omega) = \sum_i x_i \overbrace{\sum_{\substack{\omega \\ X(\omega) = x_i}}}^{p_i} = \sum_i x_i p_i$$

EX.. PRIEMER POZOROVANÍ MÁ ZMYSEL IBA VTEDY, AK POZOROVANÍ JE VEĽA.

VETA: MÁME NÁHODNÚ VELIČINU X A NEJAKÚ FUNKCIU g. POTOM $Y = g(X)$ JE TAKISTO NÁHODNÁ VELIČINA.

$$E\,g(X) = \sum_i g(x_i)\, p_i$$

$$\omega \longmapsto X(\omega) \xmapsto{g} g(X(\omega))$$

$$Eg(X) = \sum_{\omega} g(X(\omega))\, P(\omega) =$$

$$= \sum_i \sum_{\substack{\omega \\ X(\omega)=x_i}} g(\overbrace{X(\omega)}^{x_i})\, P(\omega) =$$

$$= \sum_i g(x_i) \overbrace{\sum_{\substack{\omega \\ X(\omega)=x_i}} P(\omega)}^{p_i} = \sum_i g(x_i) \cdot p_i$$

ROZPTYL $\operatorname{var} X = \underbrace{E(X-EX)^2}_{g:\ X \longmapsto (X-EX)^2} = \underbrace{EX^2}_{g^*:\ X \mapsto X^2} - (EX)^2$

$$\begin{cases} = \sum_i (x_i - EX)^2 p_i \\ = \sum_i x_i^2 p_i - \left(\sum_i x_i p_i\right)^2 \end{cases}$$

NECH X JE NÁHODNÁ VELIČINA (x_i, p_i)

VYTVÁRAJÚCA FUNKCIA $A(x) = \sum_i p_i x^i$

NAOPAK $p_i = \dfrac{A^{(i)}(0)}{i!}$

$A'(x) = \sum p_i\, i\, x^{i-1} \Rightarrow A'(1) = EX.$

05.11.2007 Mo
PETER ČERNO

MA1059 PRAVDEPODOBNOSŤ A ŠTATISTIKA

AKO SPOČÍTAŤ ROZPTYL?

$$A''(x) = \sum_i p_i \, i(i-1) \, x^{i-2}$$

$$\Rightarrow A''(1) = \sum_i i(i-1) p_i = EX(X-1)$$

$$g: X \mapsto X(X-1)$$

$$= E[X^2 - X] = EX^2 - EX$$

$$\Rightarrow \operatorname{var} X = EX^2 - (EX)^2 = EX(X-1) + EX - (EX)^2$$

VETA: $\underline{\underline{\operatorname{var} X = A''(1) + A'(1) - (A'(1))^2}}$

PRÍKLAD: $Y \sim Bi(n,p)$ (BINOMICKÉ ROZDELENIE)

$Y = X_1 + \dots + X_n$, KDE $X_i \sim Alt(p)$ NEZÁVISLÉ.

$$X_i = \begin{cases} 0 & 1-p = q \\ 1 & p \end{cases}$$

$$A_{X_i}(t) = q + p \cdot t$$

$$A_Y(t) = A_{X_1 + \dots + X_n}(t) = \underline{\underline{(q + p.t)^n}}$$

$$\operatorname{var} Y = \sum_{i=0}^{n} i^2 \binom{n}{i} p^i (1-p)^{n-i} - \left(\sum_{i=0}^{n} i \binom{n}{i} p^i (1-p)^{n-i} \right)^2 = ?$$

$$A_Y'(t) = np(q+pt)^{n-1}, \quad A_Y''(t) = n(n-1)p^2(q+pt)^{n-2}$$

$$\Rightarrow EY = A_Y'(1) = np, \quad A_Y''(1) = n(n-1)p^2$$

$$\Rightarrow \operatorname{var} Y = n(n-1)p^2 + np - (np)^2 = np(1-p) = \underline{\underline{npq}}.$$

ĎALEJ PLATÍ $E\sum X_i = \sum EX_i$ PRE X_i NEZÁVISLÉ

$$\operatorname{var}\sum X_i = \sum \operatorname{var} X_i$$

PRÍKLAD:

$$S_N = X_1 + \dots + X_N$$

$$X_i \sim Alt(p)$$

$$N \sim Pois(\lambda)$$

S_N JE NÁHODNÝ SÚČET NÁHODNÝCH VELIČÍN

$$P(S_n = k) = ?$$

X_i ... x_i, p_i $\quad \begin{cases} 0 & q \\ 1 & p \end{cases}$

Y ... y_i, q_i $\quad 0, 1, 2, \dots \quad \frac{e^{-\lambda}\lambda^i}{i!}$

S_N ... s_i, r_i $\quad 0, 1, 2, \dots \quad ?r_i?$

VETA O ÚPLNOSTI: $\Omega = \dot{\bigcup} H_i$

$$\Rightarrow P(A) = \sum_i P(A|H_i)\, P(H_i)$$

$$r_k = P(S_N = k) = \sum_{n=k}^{\infty} \overbrace{P(S_N = k \mid N = n)}^{P(S_n = k)} \cdot \underbrace{P(N = n)}_{q_n}$$

H_n .. POČET ZNESENÝCH VAJÍČOK

$$= \sum_{n=k}^{\infty} \underbrace{P(S_n = k)}_{P(\underbrace{X_1 + \dots + X_n}_{Bi(n,p)} = k)} \cdot q_n = \sum_{n=k}^{\infty} P(Bi(n,p) = k)\, q_n$$

05.11.2007 Mo
PETER ČERNO

MA1059 PRAVD. A ŠTATISTIKA

$$= \sum_{n=k}^{\infty} \binom{n}{k} p^k (1-p)^{n-k} \cdot \frac{e^{-\lambda}\lambda^n}{n!} =$$

$$= e^{-\lambda} p^k \sum_{n=k}^{\infty} \frac{\cancel{n!}}{k!(n-k)!} (1-p)^{n-k} \frac{\lambda^n}{\cancel{n!}} =$$

$$= \frac{e^{-\lambda} p^k}{k!} \lambda^k \sum_{n=k}^{\infty} \frac{1}{(n-k)!} [(1-p)\lambda]^{n-k} =$$

$$= e^{-\lambda} \frac{(\lambda p)^k}{k!} \lambda^k \sum_{n=0}^{\infty} \frac{(\lambda(1-p))^n}{n!} =$$

$$= e^{-\lambda+\lambda(1-p)} \frac{(\lambda p)^k}{k!} = \underline{\underline{e^{-\lambda p} \frac{(\lambda p)^k}{k!}}}$$

S_N SA RIADI POISSONOVÍM ROZDELENÍM, ALE <u>S PARAMETROM λp</u>.

ZÁVERY: $E\,\mathrm{Poiss}(\lambda) = \lambda$

$\Rightarrow E\,\mathrm{Poiss}(\lambda p) = \lambda p = EN \cdot EX$.

$\mathrm{var}\,\mathrm{Poiss}(\lambda) = \lambda$

$\Rightarrow \mathrm{var}\,\mathrm{Poiss}(\lambda p) = \lambda p$

OČAKÁVANÝ POČET JE $\lambda p \pm 3\sqrt{\lambda p}$ S PRAVDEPODOBNOSŤOU 99,7%.

CVIČENIE: X_i $0,1,2,\ldots A(x)$ s PRAVD. $p_0, p_1, \ldots$

N $0,1,2,\ldots B(x)$ s PRAVD: $q_0, q_1, \ldots$

$S_N = X_1 + \ldots + X_N$ $0,1,2,\ldots C(x)$ s PRAVD. $r_0, r_1, \ldots$

POTOM $C(x) = B(A(x))$.

$$E\,S_N = B'(A(1))\cdot A'(1) = B(1)\cdot A'(1) = EX_N \cdot EN$$

$$\operatorname{var} S_N = \dots$$

PREPOČÍTAŤ !

$$a > b \Rightarrow a^2 > b^2$$

ČEBYŠEVOVA NEROVNOSŤ

VETA: $P(|X - EX| \geq \varepsilon) \leq \dfrac{\operatorname{var} X}{\varepsilon^2}$

DÔKAZ: $\operatorname{var} X = \sum_i p_i (x_i - EX)^2 \geq$

$$\geq \sum_{i \mid |x_i - EX| \geq \varepsilon} p_i \underbrace{(x_i - EX)^2}_{\geq \varepsilon^2} \geq \varepsilon^2 \sum_{i \mid |x_i - EX| \geq \varepsilon} p_i =$$

$$= \varepsilon^2 P(|X - EX| \geq \varepsilon).$$

VETA: $X_1, \dots, X_n$ SÚ NEZÁVISLÉ N. V. S TÝM ISTÝM ROZDELENÍM, POTOM

$$E\,\frac{1}{n}\sum_{i=1}^{n} X_i = EX_1$$

$$\operatorname{var} \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n^2}\sum_{i=1}^{n} \operatorname{var} X_i = \frac{\operatorname{var} X_1}{n}$$

A $P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - EX_1\right| \geq \varepsilon\right) \leq \dfrac{\operatorname{var} X_1}{n\varepsilon^2} \xrightarrow[n \to \infty]{} 0$

TZV. ČEBYŠEVOVA VETA.

PODSTATNÉ! $\frac{1}{n}\sum_{i=1}^{n} X_i$ KONVERGUJE K STREDNEJ HODNOTE.

$$\overline{X_n} \to EX_1$$

HOVORÍME TOMU KONVERGENCIA V PRAVDEPODOBNOSTI.

06. 11. 2007 Tu
PETER ČERNO

EX: MAI059 PRAVDEPODOBNOST A STATISTIKA

X .. $EX = \sum x_i p_i$

$Eg(X) = \sum g(x_i) p_i$

$g(x) = e^{tx}$

$\psi(t) = E e^{tx}$

$\boxed{4.12. \quad \text{PÍSOMKA}}$ → PO CENTRÁLNU LIMITNÚ VETU

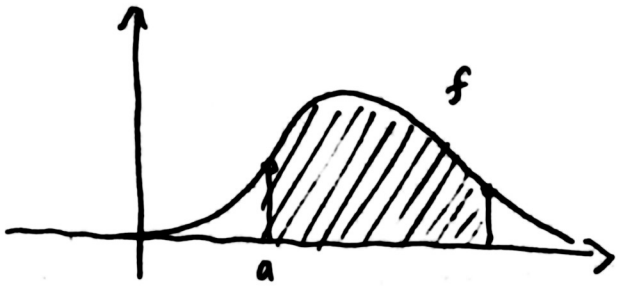
X .. NÁHODNÁ VELIČINA SO SPOJITÝM ROZDELENÍM

$P(X \leq x) = F(x) = \int_{-\infty}^{x} f(t)\,dt$

DISTRIBUČNÁ FUNKCIA

HUSTOTA

$P(x \in (a, b\rangle) = F(b) - F(a) = \int_a^b f(t)\,dt$



$P(X = a) = 0.$

$$EX = \text{VÁŽENÝ PRIEMER} = \int_{-\infty}^{\infty} \underbrace{x}_{\text{HODNOTA}} \, \underbrace{f(x)}_{\text{HUSTOTA}} \, dx$$

$$\operatorname{var} X = E(X-EX)^2 = EX^2 - (EX)^2$$

$$EX^k = \int_{-\infty}^{\infty} x^k f(x)\, dx$$

$$E(X-EX)^k = \int_{-\infty}^{\infty} (x-EX)^k f(x)\, dx$$

$$\text{ŠIKMOSŤ} = \frac{E(X-EX)^3}{[\operatorname{var} X]^{3/2}} = 0 \text{ PRE SYMETRICKÉ ROZDELENIA}$$

$$\text{ŠPICATOSŤ} = \frac{E(X-EX)^4}{[\operatorname{var} X]^{4/2}}$$

ZÁPORNÁ ~~ŠPICATOSŤ~~ ŠIKMOSŤ < 0

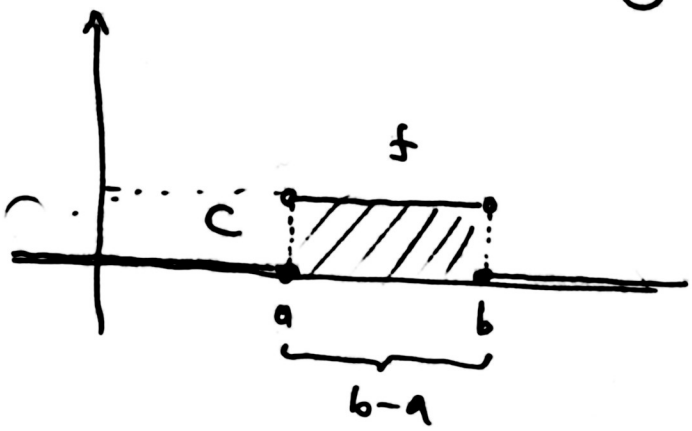
KLADNÁ ~~ŠPICATOSŤ~~ ŠIKMOSŤ > 0

06.11.2007 Tu
PETER ČERNO

EX: MA1059 PRAVD. A STATISTIKA

$X \sim Unif(a,b)$ $\quad X \in (a,b)$ S PRAVDEPOD. 1

$f(x) = const = C$ NA (a,b)
$= 0$ INDE



$$\lim_{x\to\infty} F(x) = \int_{-\infty}^{\infty} f(x)\,dx = C.(b-a) = 1$$

$$\Rightarrow C = \frac{1}{b-a}$$

$$F(X \leq \infty) = 1$$

$$EX = \int_{-\infty}^{\infty} x f(x)\,dx = \int_a^b x.C\,dx = C\left[\frac{x^2}{2}\right]_a^b =$$

$$= \frac{1}{b-a}\left(\frac{b^2}{2} - \frac{a^2}{2}\right) = \frac{a+b}{2}$$

$$E(X-EX)^2 = EX^2 - (EX)^2 =$$

$$= \left(\int_a^b x^2.C\,dx\right) - \left(\frac{a+b}{2}\right)^2 =$$

$$= \frac{1}{b-a}\cdot\frac{1}{3}\left(b^3 - a^3\right) - \left(\frac{a+b}{2}\right)^2 = \qquad \frac{(b-a)^2}{12}$$

$$= \frac{a^2+ab+b^2}{3} - \frac{a^2+2ab+b^2}{4} = \qquad \frac{-a^2-4ab-b^2}{6}$$

$$E(X-EX)^3 = 0.$$

$$\frac{1}{b-a}\int_a^b \left(x-\frac{a+b}{2}\right)^k = \frac{1}{b-a}\int_{\frac{a-b}{2}}^{\frac{b-a}{2}} u^k\,du = \frac{1}{b-a}\overset{\frac{1}{k+1}}{\left(\left(\frac{b-a}{2}\right)^{k+1}-\left(\frac{a-b}{2}\right)^{k+1}\right)}$$

$$u = x-\frac{a+b}{2}$$

$$du = dx$$

$$E(X-EX)^4 = \frac{1}{b-a}\cdot\frac{1}{5}\cdot 2\left(\frac{b-a}{2}\right)^5 =$$

$$= \frac{1}{5\cdot 2^4}\cdot(b-a)^4$$

$$E(X-EX)^k = \frac{(b-a)^k}{2^k(k+1)} \quad 2|k \qquad = 0 \text{ PRE } 2\nmid k$$

$$\text{ŠPICATOSŤ} = \frac{E(X-EX)^4}{(\operatorname{var} X)^2} = \frac{\frac{(b-a)^4}{2^4\cdot(4+1)}}{\left[\frac{(b-a)^2}{2^2\cdot(2+1)}\right]^2} =$$

$$= \frac{\frac{1}{16\cdot 5}}{\frac{1}{16\cdot 9}} = \frac{9}{5}$$

$$\text{ŠIKMOSŤ} = 0$$

06. 11. 2007 Tu
PETER ČERNO

EX: MA1059 PRAVD. A STATISTIKA

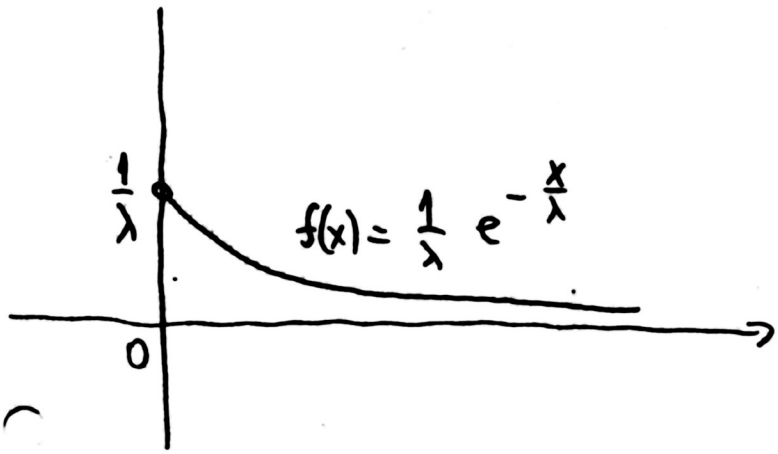
$X \sim Exp(\lambda)$ EXPONENCIÁLNE ROZDELENIE

$$f(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \quad x > 0$$
$$= 0 \quad x \leq 0$$

$\lambda > 0$ JE PARAMETER ROZDELENIA

$P(X \in (0; \infty)) = 1$

SIMULUJE DOBU MEDZI PRÍCHODMI ZÁKAZNÍKOV K FRONTE.



GAMMA FUNKCIA $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx \quad a \in \mathbb{R}$

PLATÍ $\Gamma(a+1) = a \cdot \Gamma(a)$

$\Rightarrow \Gamma(a+1) = a! \quad$ PRE $a \in \mathbb{N}$.

CHCEME SPOČÍTAŤ EX^k

$$EX^k = \int_0^\infty x^k \cdot f(x)\,dx =$$

$$= \int_0^\infty x^k \cdot \frac{1}{\lambda} e^{-\frac{x}{\lambda}}\,dx = \int_0^\infty (\lambda u)^k \cdot \frac{1}{\lambda} \cdot e^{-u} \cdot \lambda\,du =$$

$$\boxed{\begin{aligned} u &= \frac{x}{\lambda} \\ du &= \frac{dx}{\lambda} \end{aligned}}$$

$$= \lambda^k \int_0^\infty u^k e^{-u}\,du = \underline{\underline{\lambda^k \Gamma(k+1)}}$$

PRE k PRIRODZENÉ JE $\underline{\underline{EX^k = k!\,\lambda^k}}$

$$\underline{\underline{EX = \lambda}}$$

$$\operatorname{var} X = EX^2 - (EX)^2 = 2\lambda^2 - \lambda^2 = \underline{\underline{\lambda^2}}$$

$$E(X - EX)^3 = E\left(X^3 - 3X^2(EX) + 3X(EX)^2 - (EX)^3\right) =$$

$$= 6\lambda^3 - 3.2\lambda^2(\lambda) + 3.\lambda(\lambda)^2 - \lambda^3 = \underline{\underline{2\lambda^3}}$$

$$F(x) = \int_{-\infty}^x f(t)\,dt = \int_0^x f(t)\,dt =$$

$$= \int_0^x \frac{1}{\lambda} \cdot e^{-\frac{t}{\lambda}}\,dt = -\int_{-\frac{x}{\lambda}}^0 \cancel{\frac{1}{\lambda}} \cdot e^u \cdot -\cancel{\lambda}\,du =$$

$$\boxed{\begin{aligned} u &= -\frac{t}{\lambda} \\ du &= -\frac{dt}{\lambda} \end{aligned}}$$

$$= + \cancel{\frac{1}{\lambda}} \int_{-x/\lambda}^0 e^u\,du =$$

$$= + \cancel{\frac{1}{\lambda}} \left(e^0 - e^{-\frac{x}{\lambda}}\right) = \cancel{\frac{1}{\lambda}} \underline{\underline{\left(1 - e^{-\frac{x}{\lambda}}\right)}}$$

06.11.2007 Tu
PETER ČERNO

EX: MA1059 PRAVD. A STATISTIKA

STREDNÁ DOBA VAŠEHO HOVORU JE 2 min. VIETE, ŽE DĽŽKA VAŠEHO HOVORU SA RIADI EXP. ROZDELENÍM. AKÁ JE PRAVD., ŽE VÁŠ HOVOR BUDE TRVAŤ:

$$EX = \lambda = 2\,\text{min.}$$

a) PRÁVE 2 min. $\ldots = 0$

b) MENEJ NEŽ 2 min

$$P(X < 2) = 1 - e^{-\frac{2}{\lambda}} = 1 - \frac{1}{e}$$

c) NAJVIAC 2 min $= b)$

d) VIAC AKO 2 min.

$$P(X > 2) = 1 - P(X \leq 2) = e^{-\frac{2}{\lambda}} = \frac{1}{e}$$

e) 2 - 3 min.

$$P(X \in (2, 3\rangle) = F(3) - F(2) =$$

$$= \left(1 - e^{-\frac{3}{\lambda}}\right) - \left(1 - e^{-\frac{2}{\lambda}}\right) =$$

$$= e^{-\frac{2}{\lambda}} - e^{-\frac{3}{\lambda}} = \frac{1}{e} - \frac{1}{\sqrt{e^3}}$$

NÁJDITE HORNÚ HRANICU PRE DĽŽKU HOVORU TAK, ABY ĎALŠÍ HOVOR BOL S PRAVDEPODOBNOSŤOU 0,95 KRATŠÍ AKO TÁTO DĽŽKA.

$$P(X < Y) = 0{,}95 = 1 - e^{-\frac{Y}{\lambda}} \qquad e^{-\frac{Y}{\lambda}} = 0{,}05$$

$$Y = -\lambda \cdot \ln 0{,}05 = -\ln 0{,}05 \text{ min} \doteq 5{,}99 \text{ min}$$

$x = -\lambda \cdot \ln(1-p)$ $\qquad$ $F^{-1}(p)$ KVANTILOVÁ FUNKCIA

$$\ln(1-x) = 1 + \frac{x}{1} + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \frac{x^5}{5} + \dots$$

ZA DEŇ MÁME n - HOVOROV (NEZÁVISLÝCH)
URČITE ROZDELENIE (DISTR. FCIA. / HUSTOTA)

a) NAJDLHŠIEHO HOVORU ZA DEŇ.

b) NAJKRATŠIEHO HOVORU ZA DEŇ. $\qquad f(x)$

c) PRIEMERNÝ HOVOR $\overline{X}$.

b) $EY = \int_0^\infty x \cdot \left(\frac{n}{\lambda} \cdot e^{-\frac{nx}{\lambda}}\right) dx$

$u = -\frac{nx}{\lambda} \quad du = -\frac{n}{\lambda} dx$

$= \dots = \frac{\lambda}{n}$

a) MÁME $X_1, \dots, X_n$

$Y = \max(X_1, \dots, X_n)$

$$F(Y) = P(Y \le y) = \cancel{\sum_{k=0}^{n} \binom{n}{k} P(X \le y)^k \cdot P(X \le y)^{n-k}}$$

$$= P(\max X_i \le y) =$$

$$= P\left(\bigcap_{i=1}^{n} (X_i \le y)\right) = \prod_{i=1}^{n} P(X_i \le y) =$$

$$= \left(1 - e^{-\frac{y}{\lambda}}\right)^n$$

JAV X_i SÚ NEZÁVISLÉ

$\bigcup X_i \le y = X_1 \le y \vee X_2 \le y \vee \dots \vee X_n \le y$

b) $Y = \min(X_1, \dots, X_n)$

$$F(Y) = P(Y \le y) = P(\min X_i \le y) =$$

$$= P\left(\bigcup_{i=1}^{n} (X_i \le y)\right) = 1 - P(\min X_i > y) =$$

$$= 1 - P\left(\bigcap_{i=1}^{n} (X_i > y)\right) = 1 - \left[1 - \left(1 - e^{-\frac{y}{\lambda}}\right)\right]^n = 1 - e^{-\frac{ny}{\lambda}}$$

$e^{-\frac{y}{\lambda}}$

12.11.2007 Mo
PETER ČERNO

MA1059 PRAVD. A STATISTIKA

$$P(|X-EX| \geq \varepsilon) \leq \frac{\operatorname{var} X}{\varepsilon^2}$$

$$\operatorname{var} X = E(X-EX)^2$$

ČEBYŠEVOVA NEROVNOSŤ: $X_1, \dots, X_n$ NEZÁVISLÉ, S ROVNAKÝM ROZDELENÍM $EX_1 = \mu$, $\operatorname{var} X_1 = \sigma^2 < \infty$

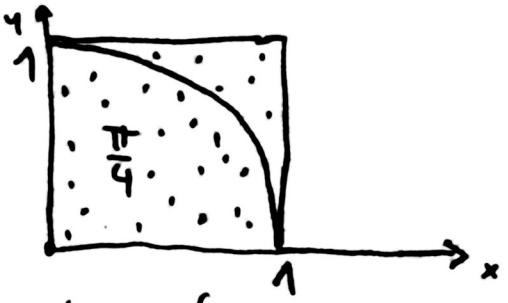
$$\bar{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i \qquad P(|\bar{X}_n - \mu| \geq \varepsilon) \leq \frac{\operatorname{var} X_1}{n\varepsilon^2} = \frac{C}{n}$$

ZREJME $E\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} EX_i = \mu$

$$\Rightarrow P(|\bar{X}_n - E\bar{X}_n| \geq \varepsilon \leq \frac{\operatorname{var}\left(\frac{1}{n}\sum X_i\right)}{\varepsilon^2} = \frac{\operatorname{var}\left(\sum_{i=1}^{n} X_i\right)}{n^2\varepsilon^2}$$

ROZPTYL BUDE KONEČNÝ PRE $\operatorname{var} X_i < C \quad \forall i$

PRÍKLAD: NAŠIM CIEĽOM JE ODHADNÚŤ π



NÁHODNÉMU BODU X_i $\begin{cases} 0 & 1-\pi/4. \\ 1 & \pi/4 \end{cases}$

ZODPOVEDÁ: TRAFIL SI?

$X_i \sim Alt\left(\frac{\pi}{4}\right) \qquad \frac{1}{n}\sum_{i=1}^{n} X_i \to \frac{\pi}{4}$ PRE $n \to \infty$

$$P\left(\left|\overline{X}_n - \frac{\pi}{4}\right| < \varepsilon\right) > 0{,}99..9 = p$$

PRESNOST STOLAHLIVOST

$$P\left(|X - EX| \leq \varepsilon\right) \geq 1 - \frac{\operatorname{var} X}{\varepsilon^2}$$

$$P\left(\left|\overline{X}_n - \frac{\pi}{4}\right| < \varepsilon\right) > 1 - \frac{\operatorname{var} \overline{X}_n}{\varepsilon^2} = 1 - \frac{\frac{\pi}{4}\left(1 - \frac{\pi}{4}\right)}{n\varepsilon^2} \geq 1 - \frac{\frac{1}{2}\cdot\frac{1}{2}}{n\varepsilon^2} =$$

$$= 1 - \frac{1}{4n\varepsilon^2}$$

CHCEME, ABY $1 - \frac{1}{4n\varepsilon^2} \geq p$

$$4n\varepsilon^2 \geq \frac{1}{1-p}$$

$$n \geq \frac{1}{4\varepsilon^2(1-p)}$$

$$\overline{X}_n = \frac{n-1}{n}\overline{X}_{n-1} + \frac{1}{n}X_n$$

EXISTUJÚ AJ SPOJITÉ NÁHODNÉ VELIČINY

$(\Omega, \mathcal{A}, P)$ $\Omega = \mathbb{R}_1$

NA $\mathbb{R}_1$ UVAŽUJEME INTERVALY $\langle a_i, b_i)$

$X : (\mathbb{R}_1, \mathcal{B}_1, P) \mapsto (\mathbb{R}_1, \mathcal{B}_1)$

BORELOVSKÉ MNOŽINY

$P(X \leq x) = F(x)$ DISTRIBUČNÁ FUNKCIA

12.11.2007 Mo
PETER ČERNO

MAI059 PRAVD. A ŠTATISTIKA

$F(x)$ JE NEKLESAJÚCA, $\lim_{x\to-\infty} F(x)=0$, $\lim_{x\to+\infty} F(x)=1$

$F(x)$ JE SPRAVA SPOJITÁ.

PREDPOKLADAJME, ŽE EXISTUJE $F'(x) = f(x)$.

$f(x)$ JE HUSTOTA.

DEF. HUSTOTOU MÔŽE BYŤ ĽUBOVOĽNÁ NEZÁP. FCIA f TAKÁ, ŽE $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

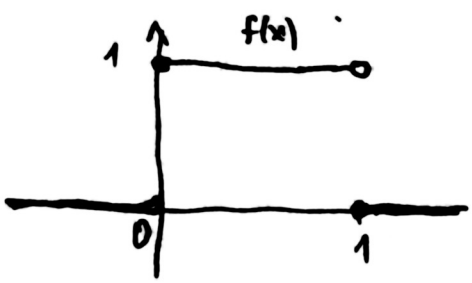
ZÁKLADNÉ CHARAKTERISTIKY

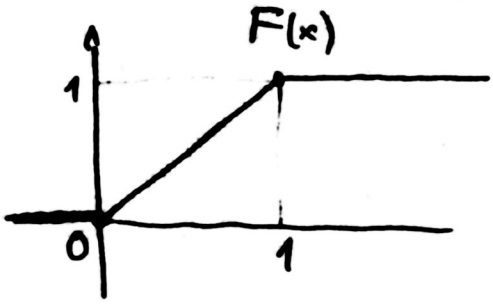
$$EX = \int_{-\infty}^{\infty} x f(x)\,dx$$

$$E_g(X) = \int_{-\infty}^{\infty} g(x) f(x)\,dx$$

$$\operatorname{var} X = \left(\int_{-\infty}^{\infty} x^2 f(x)\,dx\right) - \left(\int_{-\infty}^{\infty} x f(x)\,dx\right)^2$$

ROVNOMERNÉ ROZDELENIE NA $\langle 0,1\rangle$





VETA: $X \sim F(x)$, $F^{-1}(t)$, $R \sim R(0,1)$

PAK $Z = F^{-1}(R)$ MÁ DISTR. FUNKCIU $F(x)$

DOKAZ: $P(Z \leq x) = P(F^{-1}(R) \leq x) =$

$= P(F(F^{-1}(R)) \leq F(x)) =$

$= P(R \leq F(x)) = F(x).$ ∎

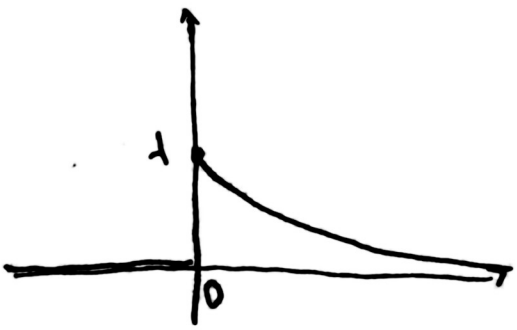
ALGORITMUS GENEROVANIA N. V. 2:

NAGENERUJ $R \sim R(0,1)$

VRÁŤ $F^{-1}(R)$

EXPONENCIÁLNE ROZDELENIE $Exp(\lambda)$

$$f(x) = \begin{cases} 0 & \text{INAK} \\ \lambda e^{-\lambda x} & x \geq 0 \quad \lambda > 0 \end{cases}$$



$EX = \frac{1}{\lambda}$

$\text{var}\, X = \frac{1}{\lambda^2}$

POUŽÍVA SA V SPOLAHLIVOSTI, FRONTÁCH ...

FUNKCIU $1-F(x)$ SA ČASTO HOVORÍ PRAVDEPODOBNOSŤ PREŽITIA

12.11.2007 Mo
PETER ČERNO

MAI057 PRAVD. A STATISTIKA

VETA: $P(X > x+y \mid X > x) = P(X > y)$

EXPONENCIÁLNE ROZDELENIE JE JEDINÉ ROZDELENIE S TOUTO VLASTNOSŤOU

$$\frac{P(X > x+y \wedge \cancel{X > x})}{P(X > x)} = \frac{1-(1+e^{-(x+y)})}{1-(1+e^{-x})} =$$

$$= e^{-y} = P(X > y).$$

VETA: $P(x < X \leq x+\Delta \mid X > x) =$

$= P($ OKAMŽITÉHO ZOBLYHANIA $) =$

$$= \frac{P(x < X \leq x+\Delta \wedge \cancel{X > x})}{P(X > x)} =$$

$$\approx \Delta \frac{f(x)}{1-F(x)} \equiv \text{INTENZITA PORUCH}$$

PRE $Exp(\lambda)$ JE $\frac{f(x)}{1-F(x)} = \lambda$.

CVIČENIE: 1) DOKÁŽTE, ŽE GEOMETRICKÉ ROZDELENIE NEMÁ PAMÄŤ

2) VHODNOU VOLBOU λ DOSTANETE Z $Exp(\lambda)$ ROZDELENIE GEOMETRICKÉ

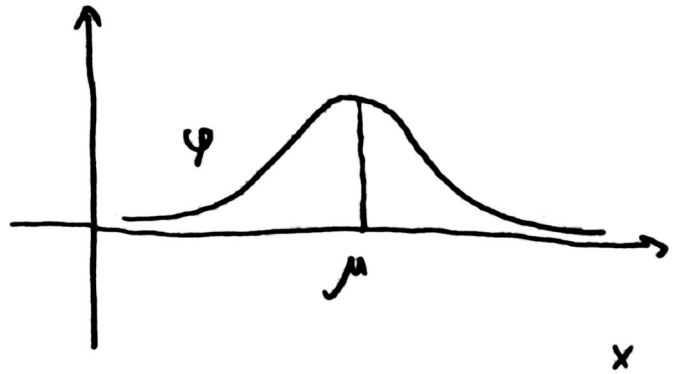
13. 11. 2007 Tu
PETER ČERNO

EX: MA1059 PRAVD. A ŠTATISTIKA

NORMÁLNE ROZDELENIE

$N(\mu, \sigma)$

$$X \to \boxed{\varphi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}}$$



PRE $N(0,1)$: $\Phi(x) = \int_{-\infty}^{x} \varphi(t)\,dt$

NEJDE SPOČÍTAŤ ANALYTICKY

Φ^{-1} KVANTILOVÁ FUNKCIA

PRÍKLAD

$N(0,1)$ $\Phi^{-1}(0{,}975) \doteq 2$ $1{,}96\ldots$

$X \sim N(0,1)$

$P(X < 2) = 0{,}975$

$P(X \geq 2) = 0{,}025$

$$P(|X| > 2) = P(X > 2) + P(X < -2) =$$
$$= 0{,}025 + 0{,}025 = 0{,}05$$

PRÍKLAD:

$Z \sim N(0,1)$

$X \sim N(\mu, \sigma^2)$ $\qquad$ $\underline{\underline{X = \mu + \sigma Z}}$

$EZ = 0$

$\operatorname{var} Z = 1$

$EX = E(\mu + \sigma Z) = \underline{\underline{\mu}}$

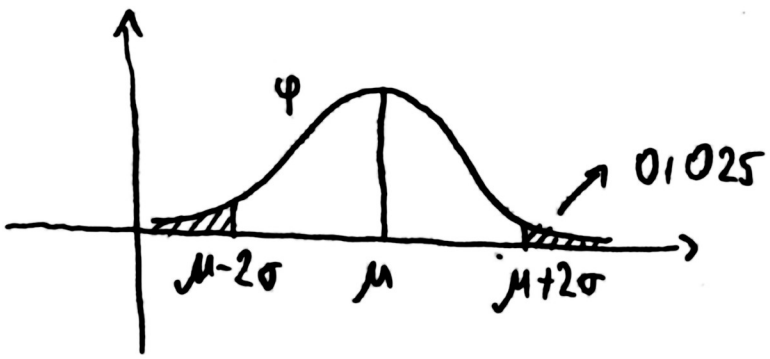
$\operatorname{var} X = \operatorname{var}(\mu + \sigma Z) = \operatorname{var}(\sigma Z) =$

$= \sigma^2 \operatorname{var}(Z) = \underline{\underline{\sigma^2}}$

$P(X > \mu + 2\sigma) = P(\mu + \sigma Z > \mu + 2\sigma) =$

$= P(Z > 2) \doteq \underline{\underline{0{,}025}}$

$P(|X - \mu| > 2\sigma) = P(|\sigma Z| > 2\sigma) = \underline{\underline{0{,}05}}$



$\Phi^{-1}(0{,}95) = 1{,}64$

13.11.2007 Tu
PETER ČERNO

EX: MAI 059 PRAVD. A ŠTATISTIKA

PRÍKLAD: IQ POPULÁCIE $\sim N(100, 15^2)$

a) AKÁ JE PRAVDEPODOBNOSŤ, ŽE V DAVE NARAZÍTE NA ADEPTA MENZY $(IQ \geq 130)$

b) NA DEBILA $(IQ \leq 70)$

$Z \sim N(0,1^2)$ $\quad IQ = \mu + \sigma Z = 100 + 15 \cdot Z$

a) $P(IQ > 130) = 1 - P(IQ \leq 130) =$

$= 1 - P(\mu + \sigma Z \leq 130) =$

$= 1 - P(100 + 15Z \leq 130) =$

$= 1 - P(Z \leq 2) = \underline{\underline{0{,}025}}$

b) $P(IQ \leq 70) = P(100 + 15Z \leq 70) =$

$= P(Z \leq -2) = \underline{\underline{0{,}025}}$

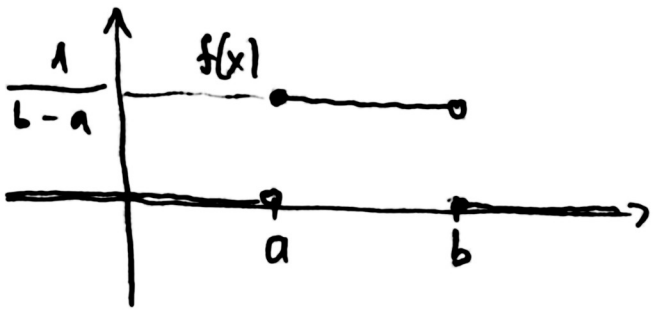
PRÍKLAD: STROJ NA VÝROBU ČOKOLÁDY VYRÁBA ŠTVORCOVÉ TABUĽKY ČOKOLÁDY. DĹŽKA STRANY TABUĽKY SA RIADI ~~NORMÁLNYM~~ ROVNOMERNÝM ROZDELENÍM NA INTERVALE (4,5) cm.

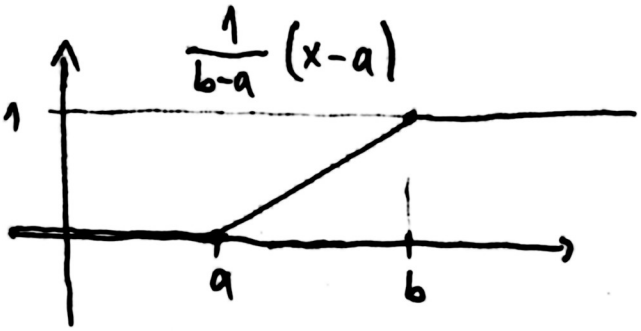
a) ROZDELENIE OBSAHU ČOKOLÁDY

b) STREDNÁ HODNOTA OBSAHU ČOKOLÁDY

$X \sim$ ROWNOMIERNIE ROZDELENIE (a, b) $\quad 0 < a < b$

$$f(x) = \begin{cases} 0 & x < a \vee x \geq b \\ \frac{1}{b-a} & a \leq x < b \end{cases} \qquad EX = \int_{-\infty}^{\infty} x f(x)\,dx$$





$Y = X^2 \qquad EX^2 = (EX)^2 + \operatorname{var} X$

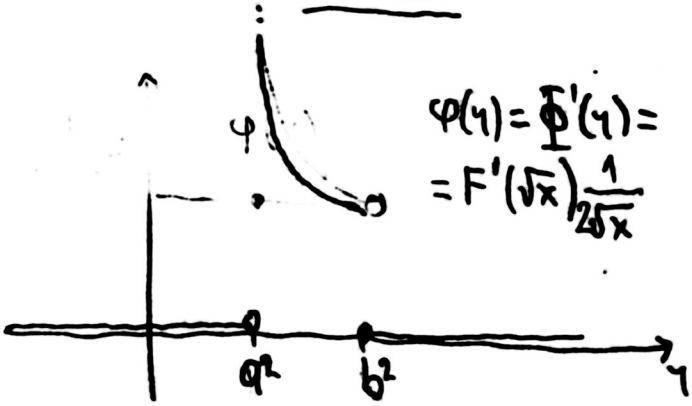
$$EX = \int_a^b x \frac{1}{b-a}\,dx = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b = \frac{a+b}{2}$$

$$EX^2 = \int_a^b x^2 \frac{1}{b-a}\,dx = \frac{1}{b-a} \cdot \frac{1}{3}(b-a)(a^2+ab+b^2) =$$

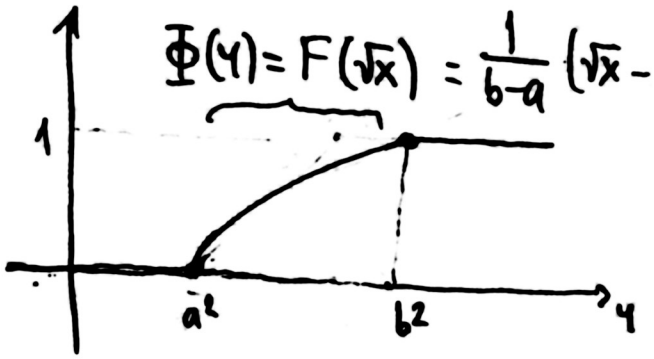
$$= \underline{\underline{\frac{a^2+ab+b^2}{3}}}$$

$$\Phi(y) = P(Y \leq y) = P(X^2 \leq x) = P(X \leq \sqrt{x}) =$$

$$= \underline{\underline{F(\sqrt{x})}}$$



$$\varphi(y) = \Phi'(y) = F'(\sqrt{x})\frac{1}{2\sqrt{x}}$$



13.11.2007 Tu
PETER ČERNO

EX: MA1059 PRAVD. A ŠTATISTIKA

$Z = (X, Y)^T$, KDE X, Y SÚ NÁHODNÉ VELIČINY

ROZDELENIE POPISUJEME DISTR. FCIOU.

$F_Z : \mathbb{R}^2 \to \langle 0, 1 \rangle$

PRE $z = (x, y)$ JE $F_Z(z) = F_Z(x, y) =$

$\underline{\underline{P(X \leq x \wedge Y \leq y)}}$. $\to$ $\boxed{\text{ZDRUŽENÁ DISTRIBUČNÁ FUNKCIA}}$

PRE DISKRÉTNE ROZDELENIA X, Y

$X \in \{x_1, \dots, x_n\}$, $Y \in \{y_1, \dots, y_m\}$

$$F_Z(x_i, y_j) = P(X \leq x_i \wedge Y \leq y_j) =$$

$$= \sum_{i : x_i \leq x} \sum_{j : y_j \leq y} P(X = x_i \wedge Y = y_j)$$

$$EZ = \begin{pmatrix} EX \\ EY \end{pmatrix}$$

$$\operatorname{var} Z = \begin{pmatrix} \operatorname{var} X & \operatorname{cov}(X, Y) \\ \operatorname{cov}(X, Y) & \operatorname{var} Y \end{pmatrix}$$

$$\operatorname{var} X = E(X - EX)^2 = EX^2 - (EX)^2$$

$$\operatorname{cov}(X, Y) = E(X - EX)(Y - EY) = \underline{\underline{E(X \cdot Y) - (EX) \cdot (EY)}}$$

$$EX \cdot Y = \sum_i \sum_j x_i y_j P(X = x_i \wedge Y = y_j)$$

MARGINÁLNA DISTR. FUNKCIA:

$$F_X(x) = P(X \leq x) = P(X \leq x \wedge Y \in \text{OBOR HODNÔT } Y) =$$

$$= \underline{\underline{F_Z(x, \infty)}} =$$

$$= \sum_{i:\, x_i \leq x} \sum_j P(X = x_i \wedge Y = y_j).$$

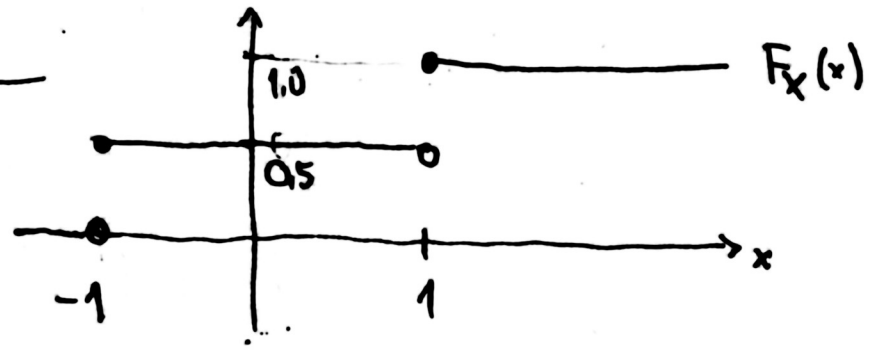
PRÍKLAD:

	Y 1	2	4	$P(X=x)$
X -1	0,1	0,25	0,15	0,5
1	0,1	0,25	0,15	0,5
$P(Y=y)$	0,2	0,5	0,3	$\Sigma = 1$

$$\underline{\underline{P(X = -1) = 0{,}5}}$$
$$\underline{\underline{P(X = +1) = 0{,}5}}$$

$$F_X(x) = P(X \leq x) = \sum_{i:\, x_i \leq x} \sum_j P(X = x_i \wedge Y = y_j).$$

X	-1	1
$F_X(x)$	0,5	1,0



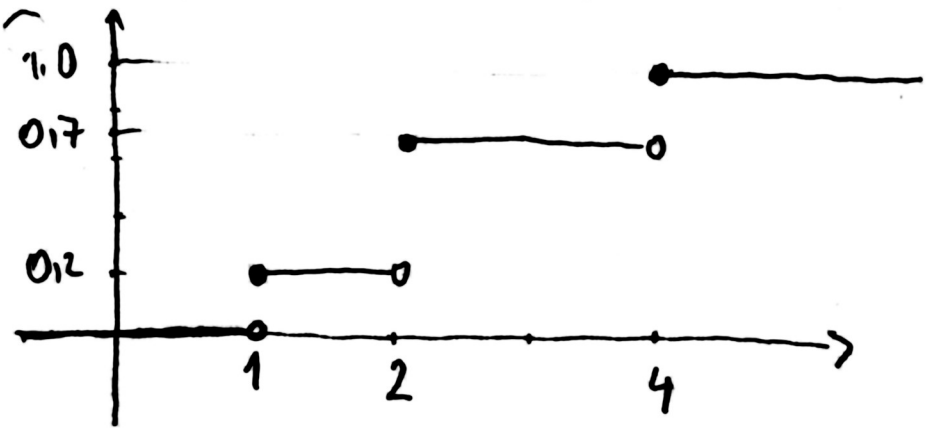
13.11.2007 Tu
PETER ČERNO

EX: MA1059 PRAVDEPODOBNOSŤ A ŠTATISTIKA

$$F_Y(y) = P(Y \le y) = \sum_i \sum_{j: y_j \le y} P(X = x_i \wedge Y = y_j)$$

Y	1	2	4
$F_Y(y)$	0,2	0,7	1,0

$P(Y=1) = 0{,}2$
$P(Y=2) = 0{,}5$
$P(Y=4) = 0{,}3$



$F_Y(y)$

$EX = 0$

$EY = 0{,}2 + 1 + 1{,}2 = 2{,}4$

$$EX \cdot Y = \sum_i \sum_j x_i \cdot x_j \cdot P(X = x_i, Y = y_j) =$$

$$= -1 \cdot 1 \cdot 0{,}1 + -1 \cdot 2 \cdot 0{,}25 + -1 \cdot 4 \cdot 0{,}15 +$$
$$+ 1 \cdot 1 \cdot 0{,}1 + 1 \cdot 2 \cdot 0{,}25 + 1 \cdot 4 \cdot 0{,}15 = \underline{\underline{0}}$$

$$\operatorname{cov}(X, Y) = E(X \cdot Y) - (EX) \cdot (EY) = \underline{\underline{0}}$$

NÁHODNÉ VELIČINY X, Y SÚ NEZÁVISLÉ,

AK $F_Z(X,Y) = F_X(X) \cdot F_Y(Y)$

PRE DISKRÉTNE ROZDELENIE:

$$\forall (i,j) \quad P(X = x_i \wedge Y = y_j) = P(X = x_i) \cdot P(Y = y_j)$$

AK SÚ X, Y NEZÁVISLÉ,
POTOM TIEŽ PLATÍ

$$E(X \cdot Y) = (EX)(EY)$$

PRE SPOJITÝ NÁHODNÝ VEKTOR

$$Z = (X,Y)^T, \quad F_Z(x,y) = P(X \leq x, Y \leq y) =$$

$$= \int_{-\infty}^{x} \int_{-\infty}^{y} f_Z(u,v)\, dv\, du$$

$$F_X(x) = P(X \leq x) = P(X \leq x \wedge Y \in \mathbb{R}) =$$

$$= \int_{-\infty}^{x} \underbrace{\int_{-\infty}^{\infty} f_Z(u,v)\, dv}_{f_X(u)}\, du$$

$$EX = \int_{-\infty}^{\infty} x\, f_X(x)\, dx$$

$$EX \cdot Y = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \cdot y\, f_Z(x,y)\, dy\, dx$$

X, Y SÚ NEZÁVISLÉ

$$\Leftrightarrow \quad f_Z(x,y) = f_X(x) \cdot f_Y(y)$$

13.11.2007 TU
PETER ČERNO

EX: MAI05Y PRAVD. A STATISTIKA

$X \sim Unif(a,b) \qquad 0 < a < b$

$Y = X^2$ (OBSAH ČOKOLÁDY)

PRE $a=4,\ b=5$

JE $EX = \frac{9}{2}$, $EY = \frac{61}{3}$

$$cov(X,Y) = E(X \cdot Y) - (EX)\cdot(EY) = \frac{3}{4}$$

$$EX^k = \frac{b^{k+1} - a^{k+1}}{(b-a)(k+1)} \quad \text{PRE } Unif(a,b)$$

KORELÁCIA:

$$\underline{\underline{Corr(X,Y)}} = \underline{\underline{\frac{cov(X,Y)}{\sqrt{var(X)\cdot var(Y)}}}} \in \langle -1, 1 \rangle$$

PRE NEZÁVISLÉ DÁTA JE VŽDY $corr(X,Y) = 0$.

PRE LINEÁRNE ZÁVISLOSTI MEDZI X A Y PLATÍ $|corr(X,Y)| = 1$.

$$corr(X,Y) = \frac{3}{4}\cdot\sqrt{\frac{135}{76}} \doteq \underline{\underline{0{,}99958}}$$

20.11.2007 Tu
PETER ČERNO

EX: MAI059 PRAVD. A ŠTATISTIKA

ČEBYŠEVOVA NEROVNOSŤ

X NAH. V. $EX = \mu$, $\operatorname{var} X = \sigma^2 < \infty$

$\forall \varepsilon > 0 \quad P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$

PRÍKLAD: $Y_i = \begin{cases} 0 & 1-p \\ 1 & p \end{cases}$ $\quad Y_1, \dots, Y_n$ SÚ NEZÁVISLÉ

$X = \sum_{i=1}^{n} Y_i$, $\quad EX = np$ $\quad X \sim Bin(n, p)$

$\operatorname{var} X = n \cdot p(1-p)$

$EY_i = p$

$\operatorname{var}(Y_i) = p(1-p)$

$E\frac{X}{n} = p$

$\operatorname{var}\left(\frac{X}{n}\right) = \frac{p(1-p)}{n}$

$$P\left(\left|\frac{X}{n} - p\right| \geq \varepsilon\right) \leq \frac{p(1-p)}{n\varepsilon^2}$$

TJ. $\lim_{n \to \infty} \frac{X}{n} = p$ (TZV. SLABÁ KONZISTENTNOSŤ)

ZÁKON VEĽKÝCH ČÍSIEL

$X_1, \dots, X_n$ N.V. PO DVOCH NEZÁVISLÉ (PO DVOCH NEKORELOVANÉ)

TJ. $\operatorname{cov}(X_i, X_j) = 0$

$EX_i = \mu$, $\operatorname{var} X_i = \sigma^2$

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right) \to 0$$

$$\leq \frac{\sigma^2}{n\varepsilon^2}$$

$$\operatorname{var}\left(\frac{1}{n}\sum X_i\right) = \frac{1}{n^2}\sum \operatorname{var} X_i = \frac{\sigma^2}{n}$$

SILNÝ ZÁKON VEĽKÝCH ČÍSIEL

$X_1, \dots, X_n$ ROVNAKO ROZDELENÉ (iid), NEZÁVISLÉ, $EX_i = \mu$

$$P\left(\frac{1}{n}\sum X_i \to \mu\right) = 1$$

CENTRÁLNA LIMITNÁ VETA:

$X_1, \dots, X_n$ NEZÁVISLÉ, $E|X_i|^3 < \infty$

$EX_i = \mu_i$, $\operatorname{var} X_i = \sigma_i^2$

$$S_n = \sum_{i=1}^{n} X_i$$

$$P\left(\underbrace{\frac{S_n - \sum_{i=1}^{n} \mu_i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}}}_{\frac{\sum (x_i - \mu_i)}{\sqrt{\sum \operatorname{var} X_i}}} \leq x\right) \longrightarrow \Phi(x)$$

20.11.2007 Ta
PETER ČERNO

EX: MA1059 PRAVD. A STATISTIKA

KEĎ $\sigma_i^2 = \sigma^2 \ \forall i$

$$P\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right) \le x\right) \to \Phi(x)$$

PRÍKLAD: STROJ NA GEN. NÁH. ČÍSIEL GENERUJE ČÍSLA 0..9 KAŽDÉ S ROVNAKOU PRAVD.

KOĽKO NAJMENEJ MUSÍME VYGENEROVAŤ ČÍSIEL, ABY SME S PRAVD. 0,975 DOSTALI ASPOŇ 2 PÁRNE ČÍSLA.

$$X_i \begin{cases} 1 & \text{i-TE VYGEN. ČÍSLO JE PÁRNE} \quad p = 0{,}5 \\ 0 & \text{i-TE VYGEN. ČÍSLO JE NEPÁRNE} \quad 1-p = 0{,}5 \end{cases}$$

$EX_i = p$

$\operatorname{var} X_i = p(1-p)$

$E|X_i|^3 < \infty$

$S_n = \sum_{i=1}^{n} X_i$ NÁH. VELIČINA, KTORÁ OZNAČUJE POČET VYGENEROVANÝCH PÁRNYCH ČÍSIEL

$$P(\underbrace{S_n}_{\sim Bin(n,p)} \ge 2) \ge 0{,}975$$

$$= \sum_{k=2}^{n}\binom{n}{k}p^k(1-p)^{n-k} = 1-(1+n)\cdot 0{,}5^n$$

$$P\left(\sum_{i=1}^{n} X_i \geq 2\right) =$$

$$= P\left(\left(\sum_{i=1}^{n} X_i\right) - n\mu \geq 2 - n\mu\right) =$$

$$= 1 - P\left(\frac{\left(\sum_{i=1}^{n} X_i\right) - n\mu}{\sqrt{n\sigma^2}} \leq \frac{2 - n\mu}{\sqrt{n\sigma^2}}\right) \doteq$$

$$\doteq 1 - \Phi\left(\frac{2 - n\mu}{\sqrt{n\sigma^2}}\right)$$

$\mu = p$

$\sigma^2 = p(1-p)$

$$1 - \Phi\left(\frac{4-n}{\sqrt{n}}\right) \geq 0{,}975$$

$$\Phi\left(\frac{4-n}{\sqrt{n}}\right) \leq 0{,}025$$

$$\frac{4-n}{\sqrt{n}} \leq -2$$

$$\sqrt{n} = \frac{2 \pm \sqrt{20}}{2} \qquad n \doteq 10{,}29$$

$\Phi^{-1}(0{,}025) \doteq -2$

$\Phi^{-1}(0{,}975) \doteq 2$

$\Phi^{-1}(0{,}05) = -1{,}64$

$\Phi^{-1}(0{,}95) = 1{,}64$

PRESNÉ PRAVDEPODOBNOSTI

$$P\left(\sum_{i=1}^{8} X_i \geq 2\right) = 0{,}965$$

$$P\left(\sum_{i=1}^{9} X_i \geq 2\right) = 0{,}980$$

20.11.2007 Ut
PETER ČERNO

EX: MA1059 PRAVD. A ŠTATISTIKA

PRÍKLAD: DOBA OBSLUHY 1 ZÁKAZNÍKA SA RIADI EXP ROZDELENÍM SO STREDNOU HODNOTOU 1 min. AKÁ JE PRAVDEPODOBNOSŤ, ŽE 20 ZÁKAZNÍKOV OBSLÚŽIM DO 20 min. ?

$X_i \sim Exp(\lambda)$, $\lambda = 1$, $EX = \lambda$, $\operatorname{var} X = \lambda^2$

$S_n = \sum_{i=1}^{n} X_i$

$$P\left(\frac{S_n - ES_n}{\sqrt{\operatorname{var} S_n}} \leq \frac{x - ES_n}{\sqrt{\operatorname{var} S_n}}\right) \doteq \Phi\left(\frac{x - ES_n}{\sqrt{\operatorname{var} S_n}}\right)$$

PRÍKLAD: ŠÍRKA OBALU $\sim N(10\,mm,\ 0{,}25\,mm^2)$
$\rightarrow$ DO KRABICE S VNÚT. ŠÍRKOU $\sim N(1014\,mm,\ 24\,mm^2)$.
AKÁ JE PRAVDEPODOBNOSŤ (PRIBLIŽNE), ŽE SA DO KRABICE NEVOJDE 100 CD-ČIEK

$X_i \sim$ ŠÍRKA i-TEHO OBALU
$Y \sim$ ŠÍRKA KRABICE

CHCEME SPOČÍTAŤ $P(\underbrace{\sum X_i - Y}_{S_n} > 0)$

$$ES_n = E\left(\sum_{i=1}^{n} X_i - 4\right) = \sum_{i=1}^{n} EX_i - E4 = -14$$

$$\operatorname{var} S_n = \sum_{i=1}^{n} \operatorname{var}(X_i) + \operatorname{var}(Y_i) = 49$$

$$P\left(\sum X_i - 4 > 0\right) =$$

$$= P\left(\frac{\sum X_i - 4 - ES_n}{\sqrt{\operatorname{var} S_n}} > \frac{-ES_n}{\sqrt{\operatorname{var} S_n}}\right) =$$

$$= 1 - P\left(\frac{S_n - ES_n}{\sqrt{\operatorname{var} S_n}} \leq \frac{-ES_n}{\sqrt{\operatorname{var} S_n}}\right) \doteq$$

$$\doteq 1 - \Phi\left(\frac{14}{\sqrt{49}}\right) = 1 - \Phi(2) = \underline{\underline{0{,}025}}$$

PRÍKLAD: ZAUJÍMA NÁS NEZNÁMY PODIEL p OSÔB, KTORÉ PIJÚ PIVO. KOĽKO OSÔB SA MÁM SPÝTAŤ, ČI PIJÚ PIVO, ABY SOM S PRAVD. ASPOŇ 0,9 ODHADOL NEZNÁMEJ p S CHYBOU NAJVIAC 0,05 ?

$$X_i = \begin{cases} 1 & p \\ 0 & 1-p \end{cases} \qquad Y = \sum_{i=1}^{n} X_i$$

$$P\left(\left|\frac{X}{n} - p\right| \geq \varepsilon\right) \leq \frac{p(1-p)}{n\varepsilon^2} \leq \frac{1}{4n\varepsilon^2}$$

$$P\left(\left|\frac{X}{n} - p\right| < \varepsilon\right) \geq 1 - \frac{1}{4n\varepsilon^2} \geq P = 0{,}9$$

$$\frac{1}{4n\varepsilon^2} \leq 1 - P \qquad \underline{\underline{n \geq \frac{1}{4\varepsilon^2(1-P)}}}$$

20.11.2007 Tu
PETER ČERNO

EX: MAI059 PRAVD. A ŠTATISTIKA

PRÍKLAD

$$X_i = \begin{cases} 1 & p \\ 0 & \end{cases}$$ i-TÝ VÝROBOK JE VADNÝ

VO VEĽKEJ ZÁSIELKE JE 10 % ZLÝCH VÝROBKOV.

JA SKONTROLUJEM n VÝROBKOV A ZÁSIELKU ODMIETNEM, POKIAĽ JE 7% ALEBO VIAC VADNÝCH VÝROBKOV.

KOĽKO MUSÍM SKONTROLOVAŤ VÝROBKOV, ABY SOM ZÁSIELKU S PRAVD. ASPOŇ 0,99 ODMIETOL?

$p = 0{,}1$ $\quad P = 0{,}99$

$q = 0{,}07$

$EX_i = p$

$\operatorname{var} X_i = p(1-p)$

$$\boxed{P\left(\sum X_i \geq qn\right) \geq p}$$

26.11.2007 Mo
PETER ČERNO

MA1059 PRAVDEPODOBNOSŤ A ŠTATISTIKA

ZAMIETNUTIE HYPOTÉZY ?

MÁME NEZÁVISLÉ JAV $X_i \sim Alt(p)$

$P(\sum X_i > C \mid \text{HYPOTÉZA})$ JE MALÁ

		MINCA V SKUTOČNOSŤ	
		✓	X
NAŠE ROZHODNUTIE	✓	OK ROZHODNUTIE	CHYBA 2. DRUHU ZLÉ ROZHOD.
	X	CHYBA 1. DRUHU ZLÉ ROZHODN.	OK ROZHODNUTIE

CHCEME, ABY $P(\text{DOBRÉ ROZHODNUTIE})$ JE VISOKÁ
$P(\text{ZLÉ ROZHODNUTIE})$ JE NÍZKA

$\boxed{\text{DOHODA}}$ $P(\text{HYP. ZAMIETNEME} \mid \text{HYP. PLATÍ})$

NAZVEME CHYBOU PRVÉHO DRUHU.

$P(\text{HYP. NEZAMIETNEME} \mid \text{HYP. NEPLATÍ})$

NAZVEME CHYBOU DRUHÉHO DRUHU

ROZHODOVACIE PRAVIDLO SA URČÍ ZA FIXOVANEJ CHYBY 1. DRUHU,

JEHO KVALITA SA OVERÍ DOPLNKOM CHYBY 2. DRUHU

$$P\left(\sum X_i > C \mid p = \tfrac{1}{2}\right) = \alpha$$

ZAFIXUJEME MALÉ PEVNÉ ČÍSLO

$$Y_n = \sum_{i=1}^{n} X_i \sim Bi(n, p)$$

$$EY_n = np, \quad var\, Y_n = np(1-p)$$

SCHEMATICKY: $P\left(Bi(n,p) > C \mid \underbrace{p = \tfrac{1}{2}}\right) = \alpha$

AKO SA ZBAVIŤ PODMIENKY?

$$\Rightarrow \quad P\left(Bi(n, \tfrac{1}{2}) > C\right) = \alpha$$

$$1 - \underbrace{P\left(Bi(n, \tfrac{1}{2}) \leq C\right)}_{\text{DISTRIBUČNÁ FUNKCIA BINOMICKÉHO ROZDELENIA}} = \alpha$$

HĽADÁM C TAK, ABY DISTR. FUNKCIA F ROZDELENIA $Bi(n, \frac{1}{2})$ NADOBUDLA HODN. $1-\alpha$.

$$F(C) = \sum_{i=0}^{C} \binom{n}{i} \left(\tfrac{1}{2}\right)^n = 1 - \alpha$$

KEĎ 100x ZOPAKUJEME TEN ISTÝ EXPERIMENT, POTOM 100α KRÁT ROZHODNEME SPRÁVNE.

26.11.2007 Mb
PETER ČERNO

MA1059 PRAVD. A STATISTIKA

$$P\left(\sum X_i > C \mid p=\tfrac{1}{2}\right) = \alpha$$

$$P\left(\frac{\sum X_i - np}{\sqrt{np(1-p)}} \leq x\right) \approx \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt$$

TJ.

$$1 - P\left(N(0,1) \leq \frac{C - 1000\cdot\frac{1}{2}}{\sqrt{1000\cdot\frac{1}{2}(1-\frac{1}{2})}}\right) = \alpha$$

$$\Rightarrow C \approx 500 + \sqrt{250}\, F^{-1}(1-\alpha)$$

PRE $\alpha = 0{,}025$ JE $F^{-1}(1-\alpha) \doteq 1{,}96$

$$C \approx np + \sqrt{np(1-p)} \cdot F^{-1}(1-\alpha)$$

$$EY_n + \sqrt{\operatorname{var} Y_n} \cdot F^{-1}(1-\alpha)$$

PRE $P\left(\sum X_i < D \mid p=\tfrac{1}{2}\right) = \alpha$

VÍDE: $D \approx np - \sqrt{np(1-p)} \cdot F^{-1}(1-\alpha)$

HYPOTÉZA $p \neq \frac{1}{2}$:

$$\underbrace{P\left(\sum X_i > C \mid p=\tfrac{1}{2}\right)}_{\frac{\alpha}{2},\ \text{POPR. } \alpha_1} + \underbrace{P\left(\sum X_i < D \mid p=\tfrac{1}{2}\right)}_{\frac{\alpha}{2},\ \text{POPR. } \alpha_2} = \alpha$$

POZNÁMKA: VÄČŠINOU SA VOLÍ $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$.

DRUHÁ CHYBA: $P(\text{NEZAMIETNEME} \mid \text{HYP. NEPLATÍ})$

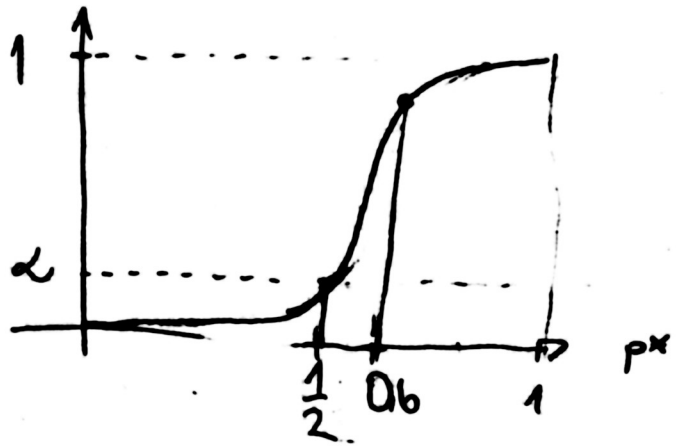
$1 - P(\ldots) = P(\text{ZAMIETNEME} \mid \text{HYP. NEPLATÍ})$

PRE NAŠU MINCU $p \in (0,1)$.

HYPOTÉZA: $p = \frac{1}{2}$

ALTERNATÍVA: $p > \frac{1}{2}$

$$\beta(p^*) = P(\text{ZAMIETNEME} \mid p = p^*) \quad \text{PRE } p^* \in (0,1)$$



ŠPECIFICKÝ PRÍPAD $p^* = \frac{1}{2}$ NÁM POSLUŽIL PRE NÁJDENIE C, D

DEFINÍCIA: HODNOTU CHYBY 1. DRUHU NAZVEME HLADINOU TESTU.

$$P\left(\frac{Y_n - np^*}{\sqrt{np^*(1-p^*)}} > \frac{C_\alpha - np^*}{\sqrt{np^*(1-p^*)}}\right) =$$

SPOČÍTANÉ K HYPOTÉZE S DANÝM α

$$= 1 - \Phi\left(\frac{532 - 1000 \cdot p^*}{\sqrt{1000 \cdot p^*(1-p^*)}}\right)$$

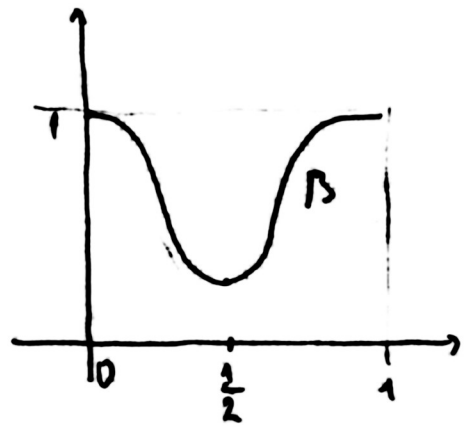
ZAUJÍMA NÁS TVAR FUNKCIE β.

26.11.2007 Mo
PETER ČERNO

MA1059 PRAVDEPODOBNOSŤ A ŠTATISTIKA

PRE HYPOTÉZU $p = \frac{1}{2}$
ALTERNATÍVU $p \neq \frac{1}{2}$

BY SME CHCELI DOSTAŤ GRAF PRE β AKO:



PRÍKLAD:
PRE $\sim N(70, 20^2)$ HMOTNOSŤ ČLOVEKA

$(10, 130)$ ... 99,7%
$(30, 110)$ ... 95%
$(50, 90)$ ... 67%

PRE VÝBER SÚČAJNÝCH ĽUDÍ:
$\mu = 70$ HYPOTÉZA
$\mu > 70$ ALTERNATÍVA

$X_1, \dots, X_n$

$$P\left(\frac{1}{n}\sum X_i > C \,\middle|\, \mu = 70\right) = \alpha$$

$$P\Big(\underbrace{\sum X_i}_{\sim N(n\mu,\, n\sigma^2)} > nC \,\Big|\, \mu = 70\Big) = \alpha$$

$$P\left(\frac{N(n\mu, n\sigma^2) - n\mu}{\sqrt{n\sigma^2}} > \frac{nC - n.\mu}{\sqrt{n\sigma^2}} \;\middle|\; \mu = 70\right) = \alpha$$

$$P\left(N(0,1) \leq \frac{nC - 70n}{20\sqrt{n}}\right) = 1 - \alpha$$

$$\frac{nC - 70n}{20\sqrt{n}} = \Phi^{-1}(1-\alpha)$$

$$C = 70 + \frac{20}{\sqrt{n}}\,\Phi^{-1}(1-\alpha)$$

$E\bar{X}_n = \mu$

$\operatorname{var}\bar{X}_n = \frac{\sigma^2}{n}$

$$\underline{\underline{E\bar{X}_n + \sqrt{\operatorname{var}\bar{X}_n}\;\Phi^{-1}(1-\alpha)}}$$

<u>POZNÁMKA !!</u> : ŠTATISTIK NEMÔŽE PRIJAŤ HYPOTÉZU. IBA JU MÔŽE ZAMIETNUŤ. !!! IBA HOVORÍME, ŽE HYPOTÉZU NEZAMIETAME.

26.11.2007/6
PETER ČERNO

MAI059 PRAVDEPODOBNOSŤ A ŠTATISTIKA

DEF.: P-HODNOTOU NAZVEME TAKÉ α, PRE KTORÚ BY SME POPRVÝKRAT HYPOTÉZU ZAMIETLI.

27.11.2007 Tu
PETER ČERNO

EX: MAI059 PRAVD. A STATISTIKA

PRÍKLAD: ODHADNITE PRAVDEPODOBNOSŤ, S AKOU BUDE POČET ŠESTIEK, KTORÉ PADNÚ V 1000 NEZÁVISLÝCH HODOCH IDEÁLNOU KOCKOU LEŽAŤ MEDZI 147 - 186.

$X_i \sim Alt\left(\frac{1}{6}\right) \qquad S_n = \sum_{i=1}^{1000} X_i$

HĽADÁME $P(147 \leq S_n \leq 186)$

$$P\left(\frac{147 - ES_n}{\sqrt{var\, S_n}} \leq \underbrace{\frac{S_n - ES_n}{\sqrt{var\, S_n}}}_{\sim N(0,1)} \leq \frac{186 - ES_n}{\sqrt{var\, S_n}}\right)$$

$$\boxed{\begin{array}{l}\Phi(1{,}96) = 0{,}975 \\ \Phi(1{,}67) = 0{,}95\end{array}}$$

$ES_n = \frac{1}{6} \cdot 1000 = \frac{500}{3} = 166{,}\bar{6}$

$var\, S_n = \frac{5}{36}\, 1000$

$$P(\ldots) = \Phi\left(\underbrace{\frac{186 - 166{,}\bar{6}}{\frac{1}{6}\sqrt{5000}}}_{\doteq 1{,}64}\right) - \Phi\left(\underbrace{\frac{147 - 166{,}\bar{6}}{\frac{1}{6}\sqrt{5000}}}_{\doteq 1{,}67}\right)$$

$$\doteq 0{,}95 - 0{,}05 = \underline{\underline{0{,}9}}$$

PRÍKLAD: ZAUJÍMA NÁS NEZNÁMY PODIEL OSÔB, KTORÉ PIJÚ PIVO. KOĽKO OSÔB SA MÁM SPÝTAŤ, ABY SOM S PRAVDEPODOBNOSŤOU ASPOŇ 0,9 ODHADLI NEZNÁME p S CHYBOU 0,05.

ČEBYŠEVOVA NEROVNOSŤ $\rightarrow$ $n \geq 1000$

$$X_i \sim Alt(p)$$

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

$$ES_n = pn$$

$$\operatorname{var} S_n = np(1-p)$$

$$p(1-p) \leq \frac{1}{4}$$

$$\sqrt{p(1-p)} \leq \frac{1}{2}$$

$$1/\sqrt{p(1-p)} \geq 2$$

$$P\left(\left|\frac{S_n}{n} - p\right| \leq 0{,}05\right) > 0{,}9$$

$$P\left(p - 0{,}05 \leq \frac{S_n}{n} \leq p + 0{,}05\right) > 0{,}9$$

$$P\left(-0{,}05n \leq S_n - ES_n \leq 0{,}05n\right) > 0{,}9$$

$$P\left(-\frac{0{,}05n}{\sqrt{n}\sqrt{p(1-p)}} \leq \frac{S_n - ES_n}{\sqrt{np(1-p)}} \leq \frac{0{,}05n}{\sqrt{n}\sqrt{p(1-p)}}\right) > 0{,}9$$

$$\Phi\left(\frac{0{,}05\sqrt{n}}{\sqrt{p(1-p)}}\right) - \Phi\left(-\frac{0{,}05\sqrt{n}}{\sqrt{p(1-p)}}\right) > 0{,}9$$

$$\Phi(0{,}1\sqrt{n}) - \Phi(-0{,}1\sqrt{n}) > 0{,}9$$

$$\Phi(0{,}1\sqrt{n}) > 0{,}95$$

$$\sqrt{n} > 16{,}7 \Rightarrow n > 279$$

16,7
167
1169
1002
167
278,89

27.11.2007 Tu
PETER ČERNO

EX: MA1059 PRAVD. A STATISTIKA

PRÍKLAD: RUDOLF DAZDÍ KAŽDÝ DEŇ DO A ZO ŠKOLY METROM V DOBE, KDE JE INTERVAL 4 min. V UVAŽOVANÝCH DENNÝCH OKAMŽIKOCH SÚ VŠETKY DOBY PRÍCHODOV RUDOLFA NA NÁSTUPIŠTE ROVNAKO PRAVDEPODOBNÉ. NECH X JE DOBA, POČAS KTOREJ RUDOLF ČAKÁ NA METRO.

a) AKÝM ROZDELENÍM SA RIADI X

b) SPOČÍTAJTE PRAVDEPODOBNOSŤ, S AKOU R. ZA 1. 1/4-ROKA (63 DNÍ) STRÁVI ČAKANÍM NA VLAK VIAC AKO 3 3/4 HOD. (225 min.)

a) $X \sim R(0,4)$ ROVNOMERNÉ ROZDELENIE

b) $S = \sum_{i=1}^{126} X_i$ $\quad P(S \geq 225)$

$EX = 2, \quad \operatorname{var} X = \frac{16}{12} = \frac{4}{3}$

$ES = 126 \cdot 2 = 252$

$\operatorname{var} S = 126 \cdot \frac{4}{3}$

$$P\left(\frac{S-252}{\sqrt{126 \cdot \frac{4}{3}}} \geq \underbrace{\frac{225-252}{\sqrt{126 \cdot \frac{4}{3}}}}_{-2,08}\right) \doteq$$

$$\doteq 1 - \Phi(-1,96) = \underline{\underline{0,975}}$$

PRÍKLAD: PROCESOR MÁ ŠÍRKU $X \sim N(10,4)$,
PATICA MÁ ŠÍRKU $Y \sim N(11,5)$
JEDEN CPU MÁ CENU 10 \$.
VYROBÍTE 1000 DVOJÍC CPU - PATICA.
AKÁ JE VAŠA STREDNÁ STRATA ?

$$Z = \begin{cases} 1 & \text{CPU} > \text{PATICA} \\ 0 & \end{cases}$$

STRATA $$W = \Big(\underbrace{\sum_{i=1}^{1000} Z_i}_{\sim Bi(1000,\, p)}\Big) \cdot 10\,\$$$

$$p = P(Z_i = 1) = P(X > Y) = P(X - Y > 0)$$

$$E(X-Y) = EX - EY = -1$$
$$\operatorname{var}(X-Y) = \operatorname{var} X + \operatorname{var} Y = 9$$

$$P(N(-1,9) > 0) =$$
$$= P\Big(\underbrace{\frac{N(-1,9)+1}{3}}_{N(0,1)} > \frac{1}{3}\Big) = 1 - \Phi\Big(\frac{1}{3}\Big) \doteq \underline{\underline{0{,}37}}$$

$$EW = \Big(\sum_{i=1}^{1000} EZ_i\Big) \cdot 10\,\$ = \underline{\underline{3700\,\$}}$$

27.11.2007 Fu
PETER ČERNO

EX: MAI059 PRAVD. A STATISTIKA

PRÍKLAD: POČET SPAMOV ZA DEŇ SA RIADI Poiss(25), MAZANIE 1 SPAMU TRVÁ 3s. AKÁ JE PRAVDEPODOBNOSŤ, ŽE ZA 100 DNÍ STRÁVIM MENEJ NEŽ 2h (7200 s) MAZANÍM SPAMOV.

$$\frac{7200\,s}{3\,s} = 2400 \text{ SPAMOV}$$

$X_i \sim \text{Poiss}(25)$ POČET SPAMOV ZA 1 DEŇ

$S = \sum_{i=1}^{100} X_i$ $\quad EX_i = 25 \quad \text{var } X_i = 25$

$$P(S \leq 2400) = P\left(\frac{S - ES}{\sqrt{\text{var } S}} \leq \frac{2400 - ES}{\sqrt{\text{var } S}}\right) =$$

$$= P\left(\frac{S - 2500}{50} \leq \frac{2400 - 2500}{50}\right) \doteq$$

$$\doteq \Phi(-2) \doteq \underline{\underline{0{,}025}}$$

PRÍKLAD: AKÁ JE PRAVDEPODOBNOSŤ, ŽE DNES DOSTANEM NAJVIAC 10 SPAMOV ZA PREDPOKLADU, ŽE RÁNO ICH MÁM V SCHRÁNKE UŽ 5.

$$p = P(X \le 10 \mid X \ge 5) = \frac{P(5 \le X \le 10)}{P(X \ge 5)}$$

$$P(X=k) = \frac{e^{-\lambda} \cdot \lambda^k}{k!} \qquad \lambda = 25$$

$$p = \frac{\sum_{k=5}^{10} \frac{e^{-\lambda} \cdot \lambda^k}{k!}}{1 - \sum_{k=0}^{4} \frac{e^{-\lambda} \lambda^k}{k!}} \doteq 0{,}000586$$

NA PÍSOMKU SI ZOBER:

KALKULAČKU !!

+ PREHĽAD ROZDELENÍ NA 1 A4 !!!

UKÁZKOVÁ PÍSOMKA (VIĎ VENCÁLEK)

PRERIEŠIŤ !

+ BAYESOVA VETA, VETA O ÚPLNEJ PRAVD.

+ POČÍTANIE KOMBINATORICKÝCH PRAVD.

+ NEZÁVISLOSŤ, KOVARIANCIA, ...

+ ČEBYŠEVOVA NEROVNOSŤ, CENT. LIMITNÁ VETA.

10.12.2007 Mo
PETER ČERNO

MA1059 PRAVD. A ŠTATISTIKA

DOPÍSAŤ POZNÁMKY ZO DŇA 03.12.2007 Mo

DEF.: NECH $X_1, \dots, X_n$ SÚ NEZÁVISLÉ $N(0,1)$ ROZDELENÉ NÁH. VELIČINY.
POTOM O N.V. $Y_n = \sum_{i=1}^{n} X_i^2$ HOVORÍME, ŽE SA RIADI TZV. χ^2-ROZDELENIM S n-STUPŇAMI VOĽNOSTI.
TOTO ROZDELENIE MÁ HUSTOTU

$$g(x) = \frac{1}{\Gamma(\frac{n}{2}) \cdot 2^{n/2}} e^{-\frac{x}{2}} x^{\frac{n}{2}-1} \qquad x > 0$$

$$= 0 \qquad x \leq 0$$

KDE $\Gamma(x) = \int_0^\infty e^{-t} \cdot t^{x-1} \, dt$ JE GAMMA FUNKCIA.
PLATÍ $\Gamma(n+1) = n!$ PRE $n \in \mathbb{N}$.

POZNÁMKA: 1) TOTO JE UMELÉ ROZDELENIE VYMYSLENÉ KVÔLI STATISTICKEJ INFERENCII.

$X_1, \dots, X_n \sim$ NEZ., $X(\mu, \sigma^2)$

$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad E\hat{\mu} = \mu$

$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\underbrace{X_i - \bar{X}}_{\sim N})^2$ JE DOBRÝ ODHAD ROZPTYLU

$\operatorname{var} X = E(X - EX)^2$

VETA: $\frac{(n-1)\hat{\sigma}^2}{\sigma^2}$ SA RIADI ROZDELENÍM χ^2_{n-1}

TOTIŽ X_i A $\bar{X}$ SÚ ZÁVISLÉ

PLATÍ $E\chi^2_n = n$, $\operatorname{var} \chi^2_n = 2n$

A PRE χ_n PLATÍ CENTRÁLNA LIMITNÁ VETA.

VETA: PRE $n \to \infty$ $P\left(\frac{\chi_n - n}{\sqrt{2n}} \leq x\right) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$

DEFINÍCIA: NECH $X \sim N(0,1)$ A $Y_n \sim \chi^2_n$.
NECH X, Y_n SÚ NEZÁVISLÉ.
POTOM $t_n = \frac{X}{\sqrt{\chi^2_n / n}}$.

MÁ TZV. ŠTUDENTOVÉ t-ROZDELENIE
O n-STUPŇOCH VOĽNOSTI.

POZNÁMKA: t_n MÁ SYMETRICKÚ HUSTOTU

$$t_\infty = \lim_{n\to\infty} t_n \sim N(0,1) \quad \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

t_1 JE TZV. CAUCHYOVSKÉ
ROZDELENIE $\frac{1}{\pi} \frac{1}{1+x^2}$.

DEF.: NECH $U_n \sim \chi^2_n$, $V_m \sim \chi^2_m$, U_n, V_m NEZAV.,
POTOM $F = \frac{1/n\, U_n}{1/m\, V_m}$ SA RIADI TZV. FISHER-SNEDECOROVÝM
ROZDELENÍM S PARAMETRAMI n, m.

10.12.2007 M.
PETER ČERNO

MA1059 PRAVD. A ŠTATISTIKA

PROBLÉM: SÚ DNEŠNÉ DETI V PRIEMERE VYŠŠÍ AKO ICH RODIČIA?

$X_1, \dots, X_n$ ... VÝŠKY SLEČIEN $\sim N(\mu, \sigma^2)$

HYPOTÉZA: $\mu = 171 = \mu_0$ JEDNOVÝBEROVÝ
ALTERNATÍVA: $\mu > 171$ TEST

1985 .. ŽENY V ČSFR PRIEMER. V. 171 cm
ROZPTYL 7^2 cm²

SÚČASTNOU VÝŠKOU ODHADNEME PRIEMER

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

KEDY ZAMIETNEME H? KEĎ $\bar{X}_n \gg 171$

$$\alpha = P(\bar{X}_n > C \mid \mu = \mu_0)$$ CHYBA PRVÉHO DRUHU

$$= P\Big(\underbrace{\frac{\bar{X}_n - \mu_0}{\sqrt{\frac{\sigma^2}{n}}}}_{\sim N(0,1)} > \frac{C - \mu_0}{\sqrt{\frac{\sigma^2}{n}}}\Big)$$

$$= 1 - P(\dots \leq \dots) = \alpha$$

POLOŽME $\frac{C - \mu_0}{\sigma} \sqrt{n} = u_{1-\alpha}$ $\qquad \alpha = 0{,}05$

$$\Rightarrow C = \mu_0 + \frac{\sigma}{\sqrt{n}} u_{1-\alpha}$$ $\qquad \Downarrow$

$$u_{1-\alpha} = 1{,}645$$

$\boxed{\text{NÁMIETKA}}$ VEĎ ALE ROZPTYL NEPOZNÁME !!!

NEZOSTÁVA NIČ INÉ, NEŽ HO NAHRADIŤ ODHADOM.

$X_1, \dots, X_n \sim N(\mu, \sigma^2)$ NEZÁVISLÉ

$\Rightarrow \bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

$\Rightarrow \bar{X}_n - \mu_0 \sim N\left(0, \boxed{\frac{\sigma^2}{n}}\right) \Rightarrow \dfrac{\bar{X}_n - \mu_0}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1)$

$$\frac{(n-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-1} \Rightarrow \hat{\sigma}^2 \sim \frac{\sigma^2 \cdot \chi^2_{n-1}}{n-1}$$

POTREBUJEME ŠTATISTIKU $\dfrac{\bar{X}_n - \mu_0}{\sqrt{\hat{\sigma}^2}} \sqrt{n} \sim t_{n-1}$

ZÁVER: t-ROZDELENIE SME SI ZAVIEDLI PRETO, ABY SME PRI NEZNÁMOM ROZPTYLE MOHLI K TESTOVANIU POUŽIŤ ŠTATISTIKU

$$\boxed{\sqrt{n}\,\frac{\bar{X}_n - \mu_0}{\hat{\sigma}}}$$

AKÝ JE ZÁKLADNÝ ROZDIEL PROTI σ^2 ZNÁMOM?

ODPOVEĎ: POUŽÍVAME INÉ KRITICKÉ HODNOTY.

PRE $n > 50$ MOŽEM VOLIŤ KVANTILY ŠTANDARDNÉHO NORMÁLNEHO ROZDELENIA NAMIESTO KVANTILOV t-ROZDELENIA. $\hat{\sigma}^2 = 1/_{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$

10.12.2007
PETER ČERNO

MAI059 PRAVD. A STATISTIKA

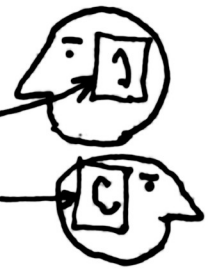
ČO MOHLO NASTAŤ:

- BUĎ SA VÝŠKA POSUNULA $N(0,1)$
- POSUNULA SA MÝŠKA A ZMENIL ROZPTYL (VARIABILITA) t_{n-1}
- ZMENIL SA ROZPTYL

ÚLOHA: SÚ UŠI ROVNAKO DLHÉ?

$l_1, \dots, l_n$ DĹŽKY ĽAVÉHO UCHA

$p_1, \dots, p_n$ DĹŽKY PRAVÉHO UCHA.



$\boxed{\text{POZOR}}$ l_i A p_i SÚ ZÁVISLÉ VELIČINY !!!

$$X_i = L_i - P_i$$

$\uparrow$

MÁME JEDNOVÝBEROVÝ PROBLÉM

$$X_i \sim N(0, \sigma^2)$$

ČOMU TO ZODPOVEDÁ?

HYPOTÉZE, ŽE UŠI SÚ ROVNAKO VEĽKÉ.

ALTERNATÍVA: $\mu > 0$.

PROBLÉM JE V TOM, ŽE DĹŽKY UŠÍ SÚ SILNO KORELOVANÉ.

11. 12. 2007 Tu
PETER ČERNO

EX: MA1059 PRAVD. A ŠTATISTIKA

ŠTATISTIKA

$x_1, x_2, \dots, x_n$ → VÝBEROVÝ SÚBOR

→ $\bar{x}$

$$s_x = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}$$

TZV. POPISNÁ ŠTATISTIKA

$x_i \leftrightarrow$ REALIZÁCIA N. V. X_i.

PRE N. V. PREDPOKLADÁME NEJAKÝ MODEL

NAPR. $X_i \sim N(\mu, \sigma^2)$

1) METÓDA MOMENTOV:

$$EX^r = \frac{1}{n} \sum x_i^r$$

PRÍKLAD: ODHAD μ, σ^2 v $N(\mu, \sigma^2)$ METÓDOU MOMENTOV.

$$\hat{\mu} = \frac{1}{n} \sum x_i = \bar{x}$$

$$\sigma^2 = EX^2 - (EX)^2$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum x_i^2 - \left(\frac{1}{n} \sum x_i\right)^2 = \frac{1}{n} \sum (x_i - \bar{x})^2$$

ČASTO KLADIEME $\hat{\sigma}^2 = \left(\frac{1}{n} \sum x_i^2\right) - \bar{x}^2$ (FUNKCIA DÁT)

$= \frac{1}{n} \sum X_i^2 - \bar{X}^2$ (FUNKCIA NÁH. VELIČÍN)

$$E\hat{\sigma}^2 = E\left(\frac{1}{n}\sum X_i^2 - \left(\frac{1}{n}\sum X_i\right)^2\right)$$

$$\left.\begin{aligned} EX &= \mu \\ \operatorname{var} X &= \frac{\sigma^2}{n} \end{aligned}\right\} \quad E\bar{X}^2 = \operatorname{var}\bar{X} + (E\bar{X})^2 = \frac{\sigma^2}{n} + \mu^2$$

$$\operatorname{var} X_i = \sigma^2 \Rightarrow EX_i^2 = \sigma^2 + \mu^2$$

$$\Rightarrow E\hat{\sigma}^2 = \frac{1}{n}\sum(\sigma^2+\mu^2) - \left(\frac{\sigma^2}{n} + \mu^2\right) = \underline{\underline{\frac{n-1}{n}\sigma^2}}$$

NÁHODNÁ VELIČINA

číslo

PRETO RADŠEJ KLADIEME $\hat{\sigma}^2 = \underline{\underline{\frac{1}{n-1}\sum(x_i - \bar{x})^2}}$

PRÍKLAD: $X_i \sim \text{Unif}(0, \Theta)$

$\hat{\Theta} = ?$

$EX_i = \frac{\Theta}{2}$

$\hat{\Theta} = \frac{2}{n}\sum x_i$

11. 12. 2007 Tu
PETER ČERNO

EX: MA1 059 PRAVD. A ŠTATISTIKA

$X_i \sim Exp(\lambda) \leftarrow$ MODEL $\quad EX_i = \lambda$

$\Rightarrow \quad \hat{\lambda} = \frac{1}{n}\sum x_i$

2) METÓDA MAXIMÁLNEJ VIEROHODNOSTI (ML)

$x_1, \dots, x_n \longleftrightarrow \underbrace{X_1, \dots, X_n}_{\text{NEZÁV., ROVN. ROZDELENÉ (iid)}}, \sim F(x_i, \Theta)$

$f(x_i, \Theta) \qquad P(X=k, \Theta)$

$\xrightarrow{\text{vieroh.}}$ $f(x_1, \dots, x_n, \Theta) = \prod f(x_i, \Theta)$

$\hat{\Theta}_{ML} = \arg\max_{\Theta} \prod f(x_i, \Theta)$

NAPR. PRI $N(\mu, \sigma^2)$ JE $\Theta = (\mu, \sigma^2)^T$

NAMIESTO $\prod f(x, \Theta)$

MAXIMALIZUJEME $\sum \log f(x, \Theta)$

LOGARITMICKÁ VIEROHODNOSŤ

PRÍKLAD: $\hat{\lambda}_{ML}$ v $Exp(\lambda)$

$f(x_i, \lambda) = \frac{1}{\lambda} e^{-\frac{x_i}{\lambda}} \qquad x_i > 0$

$\log f(x_i, \lambda) = \log\left(e^{-\frac{x_i}{\lambda}}\right) - \log(\lambda) = -\frac{x_i}{\lambda} - \log(\lambda)$

$\sum \log f(x_i, \lambda) = -n\log(\lambda) - \frac{1}{\lambda}\sum x_i = \varphi(\lambda)$

$$\varphi'(\lambda) = -\frac{n}{\lambda} + \frac{1}{\lambda^2}\left(\sum x_i\right) = 0$$

$$\Leftrightarrow \quad \lambda = \frac{1}{n}\sum x_i$$

$$\hat{\lambda}_{ML} = \bar{x}$$

3) BODOVÝ ODHAD

→ INTERVALOVÝ ODHAD

$$(\hat{\Theta}_S, \hat{\Theta}_H) \quad P_\Theta\left(\underbrace{(\hat{\Theta}_S, \hat{\Theta}_H)}_{\text{NÁH. VELIČINY}} \text{ POKRYL } \Theta\right) = 1-\alpha$$

$\hat{\Theta}$ JE NÁH. VELIČINA

POTREBUJEME POZNAŤ JEJ ROZDELENIE

$$X_i \sim N(\mu, \sigma^2)$$

$$\hat{\mu} = \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\boxed{\bar{X} \pm \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}}}$$

11.12.2007 Tu
PETER ČERNO

EX: MAI059 PRAVD. A ŠTATISTIKA

MODEL: VÝŠKA CHLAPCOV VO VEKU 9,5 - 10 ROKOV $\sim N(\mu, \sigma^2)$.

PREDPOKLADÁME, ŽE POZNÁME ROZPTYL

$\sigma^2 = 39{,}06\ cm^2$ $\qquad \sigma \doteq 6{,}25$

ZMERALI SME VÝŠKU 15 CHLAPCOV

A SPOČÍTALI $\bar{X} = 139{,}1$ cm.

a) 99% INTERVAL SPOLAHLIVOSTI PRE μ. $\underline{\underline{\alpha = 0{,}01}}$

$$\underbrace{\bar{X}}_{139{,}1} \pm \underbrace{\Phi^{-1}(0{,}995)}_{2{,}58} \frac{6{,}25}{\sqrt{15}} \qquad \underline{\underline{(134{,}9;\ 143{,}3)}}$$

b) KOĽKO NAJMENEJ JE Ø VÝŠKA CHLAPCOV?

$$\boxed{\begin{array}{l}\text{ĽAVOSTRANNÝ INTERVAL } (\Theta, \infty) \\ \bar{X} - \Phi^{-1}(1-\alpha)\frac{\sigma}{\sqrt{n}}\end{array}}$$

95% ĽAVOSTRANNÝ INT. $\alpha = 0{,}05$

$$\underbrace{\bar{X}}_{139{,}1} - \underbrace{\Phi^{-1}(0{,}95)}_{1{,}64} \frac{6{,}25}{\sqrt{15}} \doteq 136{,}4$$

KEĎ σ^2 NEPOZNÁME, POUŽÍVAME S_X^2 A t_{n-1}^{-1}

$$\bar{X} \pm t_{n-1}^{-1}\left(1 - \frac{\alpha}{2}\right) \frac{S_X}{\sqrt{n}}$$

log-VEROHODNOSŤ $= \ell(\Theta)$

$\hat{\Theta} = \operatorname{argmax} \ell(\Theta)$ $\boxed{\left(-\frac{d^2 \ell(\Theta)}{d\Theta^2}\right)^{-1}}$

DÁ SA DOKÁZAŤ, ŽE

$$\hat{\Theta}_n \sim N\left(\Theta, \left(-\frac{\partial^2 \ell(\Theta)}{\partial\Theta\,\partial\Theta^T}\right)^{-1}\right)$$

$$\hat{SE}(\hat{\Theta}) = \sqrt{\left(-\frac{d^2 \ell(\Theta)}{d\Theta^2}\right)^{-1}}$$

INT. ODHAD $\hat{\Theta} \pm \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\ldots}$

PRÍKLAD: 126 CIEST METRON

R. ČAKAL DOHROMADY 275 min.

? PRIBLIŽNÝ INTERVAL SPOLAHLIVOSTI PRE INTERVAL MEDZI METRAMI

$X_i \sim Unif(0, \theta)$

$\sum X_i$ NAPR. METÓDOU MOMENTOV

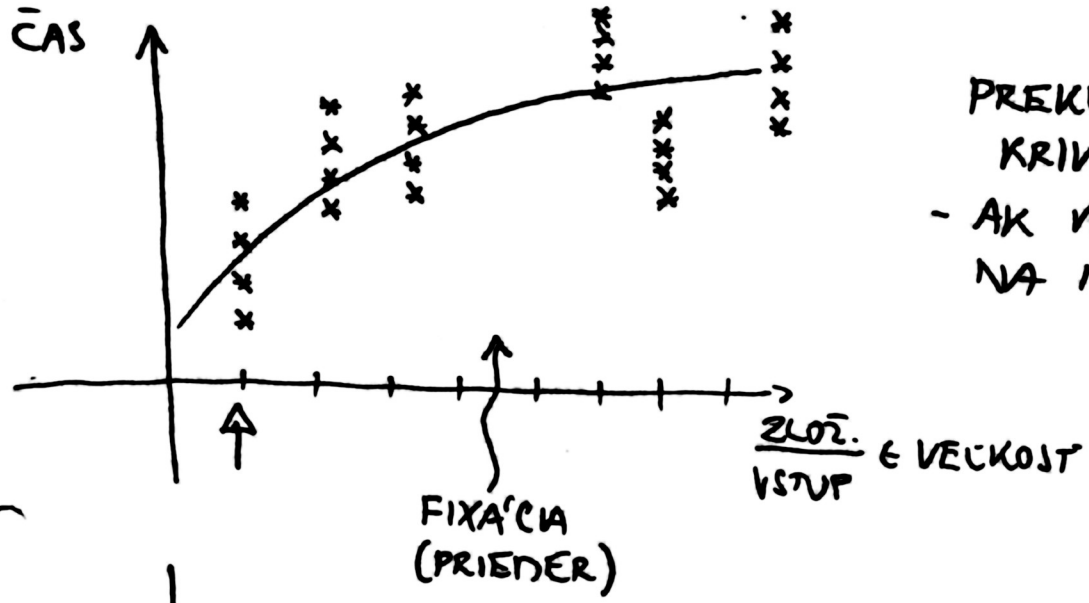
$$\hat{\Theta} = 2 \cdot \frac{\sum X_i}{n} = 2 \cdot \frac{275\,min}{126}$$

$$\frac{2\sum X_i}{n} \overset{?}{\sim} N(\ldots)$$

CENTRÁLNA LIMITNÁ VETA

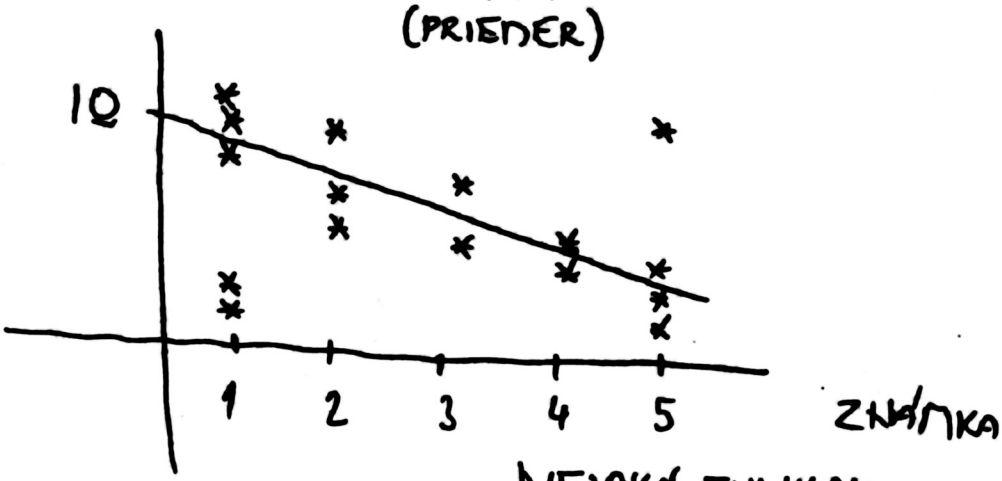
07.01.2008 Mo
PETER ČERNO

MA1059 PRAVD. A ŠTATISTIKA



PREKLADANIE KRIVIEK
- AK VEL. ZÁVISÍ NA INEJ VELIČINE

= VEĽKOSŤ



$$E(Y \mid X = x) = \mu(x)$$ = MODEL

NEJAKÁ FUNKCIA

// PREMENNÁ STREDNÁ HODNOTA

$(Y, X) \leftarrow$ ZAFIXOVANÉ X

$$E(Y \mid X = x) = \mu(x) = \beta_0 + \beta_1 x$$

n = # MERANÍ

$$Q_i = \beta_0 + \beta_1 x \qquad i = 1, 2, \dots, n$$

$Q_i = \beta_0 \rightarrow$ KONŠT. = NAJJEDNODUCHŠIE :-)

V REÁLI MERAM $(X_i, Y_i)_{i=1}^{n}$ A CHCEM ODHALIŤ ZÁVISLOSŤ.

$$IQ = \beta_0 + \beta_1 x$$ } PRIDÁM CHYBU / NÁHODU

$\uparrow$ DETERMINISTICKY

$$IQ = \beta_0 + \beta_1 x + \varepsilon$$ ← NÁHODNÁ CHYBA

CHYBY MERANIA
CHYBY MODELU
CHYBY ŠTUDENTA

Y .. ZÁVISLÁ PREMENNÁ (ODPOVEĎ)

LINEÁRNE ZÁVISÍ NA X .. VSVETLUJÚCA PREMENNÁ

A CHYBE .. NÁHODNE

↳ POSUNUTIE - SYSTEMATICKÁ CHYBA

$(X_i, Y_i)_{i=1}^{n}$ ∈ MERANIA

MODEL : VŠEOBECNÝ : $Y = \beta_0 + \beta_1 X + \varepsilon$

PRE JEDNOTLIVCOV $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$

NEMÁME SYSTEM. CHYBU

O CHYBÁCH ε_i BUDEME PREDPOKLADAŤ, ŽE SÚ NEZÁVISLÉ, ĎALEJ $\varepsilon_i \sim N(0, \sigma^2)$

SILNÝ PREDPOKLAD

β_0, β_1 - NEMÁME PARAMETRE MODELU

ε_i - NÁH. VELIČINA CHYBA NORM. ROZDELENIA

$$Y_i \sim N(\beta_0 + \beta_1 X_i, \sigma^2)$$

$$\boxed{\begin{aligned} &Z \sim N \\ &a + Z \sim N \\ &E(a+Z) = Ea + EZ = a \end{aligned}}$$

$$\bar{Y} = \beta_0 + \beta_1 \bar{X} + \bar{\varepsilon}$$

$\bar{Y}$ - PRIEMER $= \frac{1}{n} \sum_{i=1}^{n} Y_i$

$$\bar{\varepsilon} = \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i$$

$$\sim N\left(0, \frac{\sigma^2}{n}\right)$$

07.01.2008 Po
PETER ČERNO

MAI 059 PRAVD A STATISTIKA

AKO ODHADNÚŤ NEZNÁME PARAMETRE $\beta_0, \beta_1, \sigma^2$?

VIAC MERANÍ AKO PARAMETROV

ZVOLÍM STRATOVÚ FUNKCIU MINIMALIZUJÚCU:

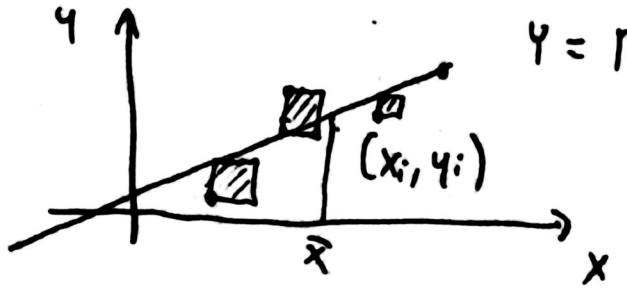
$$\sum_{i=1}^{n} \left(Y_i - \beta_0 - \beta_1 X_i \right)^2 \quad \text{II}$$

$$\sum_{i=1}^{n} \left| Y_i - \beta_0 - \beta_1 X_i \right)^2 \quad \text{I}$$

$$\underbrace{\sum_{i=1}^{n} \int \left(Y_i - \beta_0 - \beta_1 X_i \right)}_{\text{VŠEOBECNE}}$$

DEFINÍCIA: HODNOTE $r_i = Y_i - \beta_0 - \beta_1 X_i$ SA STRANDLA HOVORÍ REZIDUUM.

- CHCEME MINIMALIZOVAŤ
- GAUSS-RNS $^2 \in$ OK



$Y = \beta_0 + \beta_1 X$

GEOM.
- MIN. ŠTVORCE

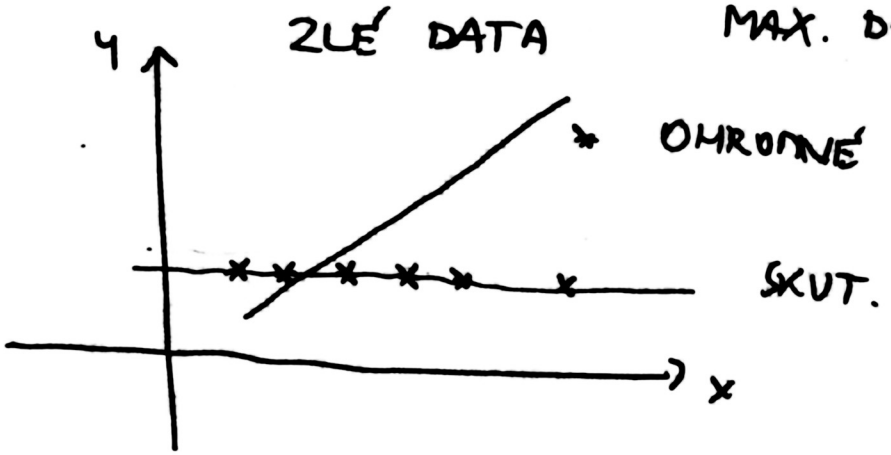
$$\left(\sum_{i=1}^{n} \left(Y_i - \beta_0 - \beta_1 X_i \right)^2 \right)'_{\beta_0} = 0$$

$$\left(\sum_{i=1}^{n} \left(Y_i - \beta_0 - \beta_1 X_i \right)^2 \right)'_{\beta_1} = 0$$

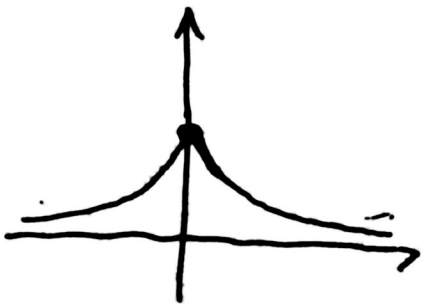
NORM. ROZD. + ... → MIN ŠTV. min $\sum r_i^2$

DVOJITÉ EXP. + 1 ... → MIN ABS ODCH. min $\sum |r_i|$

ZLÉ DATA MAX. DÔVERYHODNOSŤ



2-EXP. ROZDELENIE $\frac{1}{2\lambda} e^{-\frac{|x|}{\lambda}}$



$$\min_{\beta_0, \beta_1} \sum_{i=1}^{n} \left(Y_i - \beta_0 - \beta_1 x_i \right)^2$$

$$\left.\begin{aligned} \frac{\partial \ldots}{\partial \beta_0} &= 0 \\ \frac{\partial \ldots}{\partial \beta_1} &= 0 \end{aligned}\right\}$$

VRIEŠIM SÚSTAVU

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})^2}$$

1 ODHAD.

JE ZÍSKANÝ ODHAD „ROZUMNÝ" ?

$$E\hat{\beta}_1 = E \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} = \sum_i \frac{(X_i - \bar{X})}{\sum_j (X_j - \bar{X})^2} E(Y_j - \bar{Y})$$

$$= \beta_1 \frac{\sum (\bar{X} - X_i)^2}{\sum (\bar{X} - X_i)^2} = \beta_1$$

07.01.2008 Mo
PETER ČERNO

MAI059 PRAVD. A ŠTAT.

$$E(Y_i - \bar{Y}) = \beta_0 - \beta_1 x_i - \beta_0 - \beta_1 - \bar{x} =$$
$$= \underline{\underline{\beta_1 (X_i - \bar{X})}}$$

V PRIEMERE MÁME SPRÁVNY VÝSLEDOK

$$E\hat{\beta}_0 = \beta_0 \quad \text{(ROVNAKO)}$$
$$= E(\hat{Y} - \hat{\beta}_1 \cdot \bar{X}) = \beta_0 + \beta_1 \bar{X} - \beta_1 \bar{X} = \beta_0 .$$

PLATÍ + OVER : $\operatorname{var} \hat{\beta}_1 = \dfrac{\sigma^2}{\sum_{i=1}^{n} (X_i - \bar{X})^2}$

$$\hat{\beta}_1 = \sum_i \frac{(X_i - \bar{X})^2}{\sum_j (X_j - \bar{X})^2} \cdot \frac{(Y_j - \bar{Y})}{(X_i - \bar{X})}$$

SMERNICA

VÁŽENÝ PRIEMER SMERNÍC

AKO ODHADNÚŤ σ^2 ?

$$\min_{\beta_0, \beta_1} \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_i)^2 = \min_{\beta_0, \beta_1} \sum r_i^2$$

$$\Rightarrow \text{ODHAD} \quad \frac{\sum_{i=1}^{n} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 \bar{X})^2}{n}$$

$Y = \hat{\beta}_0 + \hat{\beta}_1 x$ ≡ ODHAD NEZNÁMEJ REGRESNEJ PRIAMKY

RSS - RESIDUAL SUM OF SQUARES

$\boxed{\text{PLATÍ:}}$ $$RSS = \sum_{i=1}^{n} (Y_i - \bar{Y})^2 - \hat{\beta}_1 \sum_{i=1}^{n} (X_i - \bar{X})^2$$

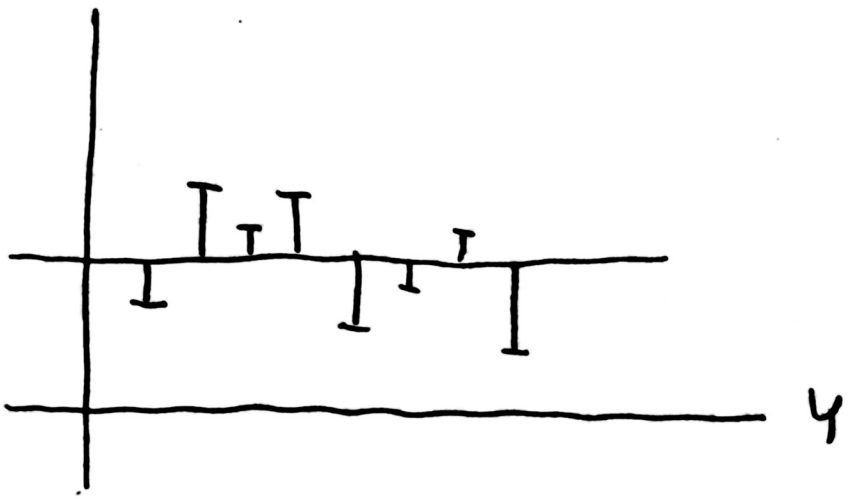
OK. ZLEPŠENIE. O KOĽKO? POUŽITÍM PRIAMKY SI POLEPŠÍM

$$\text{var} \frac{RSS}{n-2} = \sigma^2$$

POZN.: $Y = \beta_0 + \varepsilon$

– KONŠTANTY ← DÔLEŽITÉ

VÝBEROVÝ PRIEMER JE ONNČ, KT. JE OPTIMÁLNY PRE NORMÁLNE ROZDELENIE CH7B1



07.01.2008 Mo.
PETER ČERNO

MAI059 PRAVD. A STAT.

$\frac{RSS}{n-2}$ JE DOBRÝ (NESTRANNÝ) ODHAD ROZPTYLU
(2 NEZNÁME PARAMETRE)

OTÁZKA: JE VÝVOJ TEPLOTY MERANEJ V KAROLÍNE SKUTOČNE RASTÚCI?

H: VÝVOJ TEPLOTY JE KONŠT.
A: PRIEM. TEPLOTA RASTIE

H: $\beta_1 = 0$ $\qquad Y = \beta_0 + \beta_1 X + \varepsilon$
A: $\beta_1 > 0$

IQ: A: $\beta_1 < 0$ ≡ IQ KLESÁ

KEDY ZAMIETNEM HYPOTÉZU? $\boxed{\hat{\beta}_1 >> 0}$

$$\hat{\beta}_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} \sim N\left(\beta_1, \frac{\sigma^2}{\sum (X_i - \bar{X})^2}\right)$$

NECH POZNÁME σ^2

$$\Rightarrow \frac{\hat{\beta}_1 - \beta_1}{\sigma} \sqrt{\sum (X_i - \bar{X})^2} \sim N(0,1)$$

$$\alpha = P(\hat{\beta}_1 > c \mid \beta_1 = 0) =$$ CHYBA 1. DRUHU HYP. PLATÍ

σ NAHRADÍME $\hat{\sigma} = \frac{RSS}{n-2}$ A PRACUJEME S ROZDELENÍM $N(0,1) / \sqrt{\frac{\chi^2}{n}}$.

08.01.2008 Tu
PETER ČERNO

EX: MA1059 PRAVD. A ŠTATISTIKA

PRÍKLAD: ČOI NA KONTROLE V REST. NA RÚŽKU A NECHALA SI NATOČIŤ 8x 0,5l PIVA.
SKUTOČNÝ OBJEM PIVA V POHÁRI BOL:

0,510	0,462	0,491
0,466	0,461	0,503
0,495	0,488	

MÔŽE KONTROLÓR TVRDIŤ:

a) HOSTINSKÝ TOČÍ V PRIEMERE INÚ MIERU AKO 0,5l ?

b) HOSTINSKÝ ŠIDÍ HOSŤOV ?

c) HOSTINSKÝ OKRÁDA MAJITEĽA ?

$\phi = 0{,}4845 = \hat{EX}$

X = OBJEM PIVA V NÁHODNE VYBRANOM POHÁRI

CHCEME ZISTIŤ, Ó $EX \neq 0{,}5$

$H_0 = EX = 0{,}5$
$A_0 = EX \neq 0{,}5$

t - TEST "STUDENT"

$X \sim N(\mu, \underline{\underline{\sigma^2}}) \rightarrow \begin{matrix} H_a: \mu = c \\ A_a: \mu \neq c \\ \gtrless \end{matrix} \rightarrow \dfrac{\bar{X} - c}{\sqrt{\frac{\sigma^2}{n}}} \underset{\text{PRI } H_a}{\sim} N(0,1)$

AK POZNÁME σ^2

$X \sim N(\mu, \sigma^2) \qquad T = \dfrac{\bar{X} - c}{\sqrt{\frac{s^2}{n}}} \underset{\text{PRI } H_a}{\sim} t_{n-1}$

AK NEPOZNÁME σ^2

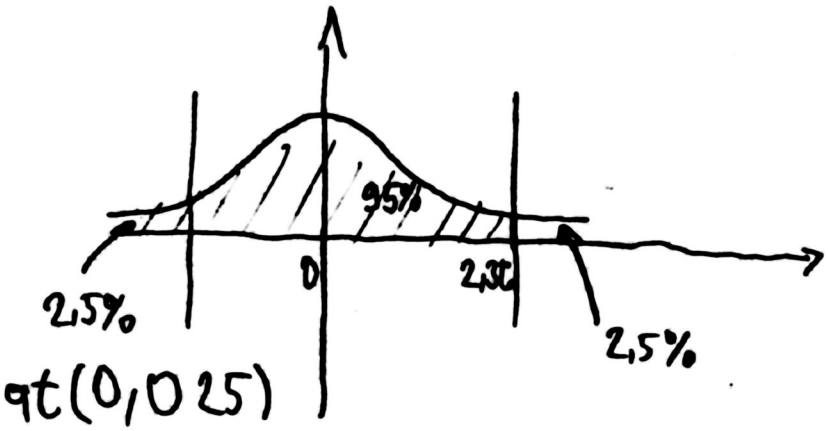
$s^2 = \dfrac{1}{n-1} \sum (x_i - \bar{x})^2$

A PRE <u>VEĽKÉ n</u> $\sim N(0,1)$

MY MÁME $n = 8$, $s = 0{,}0191$

$T = -2{,}29$

P_{H_a} (ZAMÍTNUŤ) $= \alpha$ HLADINA VÝZNAMNOSTI 0,05



P-HODNOTA = 0,055553

NEPODARILO SA NÁM UKÁZAŤ, ŽE TOČÍ NIEČO INÉ AKO 0,5

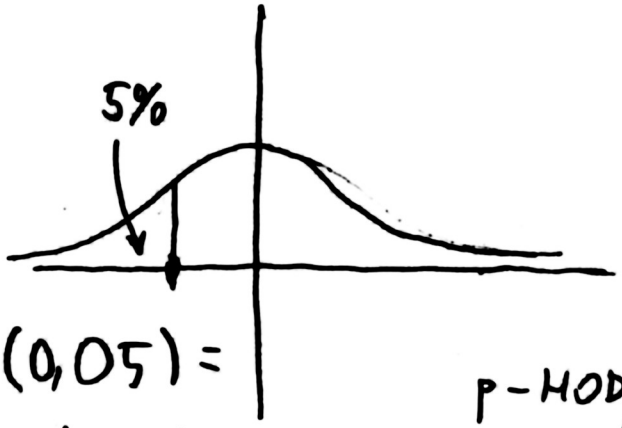
---

b) $H_b: EX = 0{,}5$

$A_b: EX < 0{,}5$

08.01.2008 Tu
PETER ČERNO

EX: MA1059 PRAVD. A STATISTIKA



$qt(0{,}05) =$

$= -qt(0{,}95) =$

$= -1{,}89$

p-HODNOTA VYJDE
0,028

PODARILO SA NÁM DOKÁZAŤ, ŽE TOČÍ MENEJ AKO 0,5 l.

c) ZREJME NEOKRÁDA MAJITEĽA

NEDOKÁZALI SME, ŽE TOČÍ VIAC AKO 0,5 l.

TO ALE NEZNAMENÁ, ŽE <u>NETOČÍ</u> VIAC AKO 0,5 l !!!

p-HODNOTA JE $1 - 0{,}028$.

PRÍKLAD: KLINICKÁ ŠTÚDIA SLEDOVALA ENERGETICKÝ PRÍJEM ANOREKTICKÝ PRÍJEM PRED A PO MENSTRUÁCII:

$n = 11$

PRED	PO
5260	3910
5470	4220
5640	3885
6180	5160
⋮	⋮
8770	7335

MÁ MENSTRUÁCIA VPLYV NA ENERG. PRÍJEM

TZV. <u>PÁROVÝ T-TEST</u>

$X = \text{PRED} - \text{PO}$

$\phi = \widehat{EX}$

$H: EX = 0$

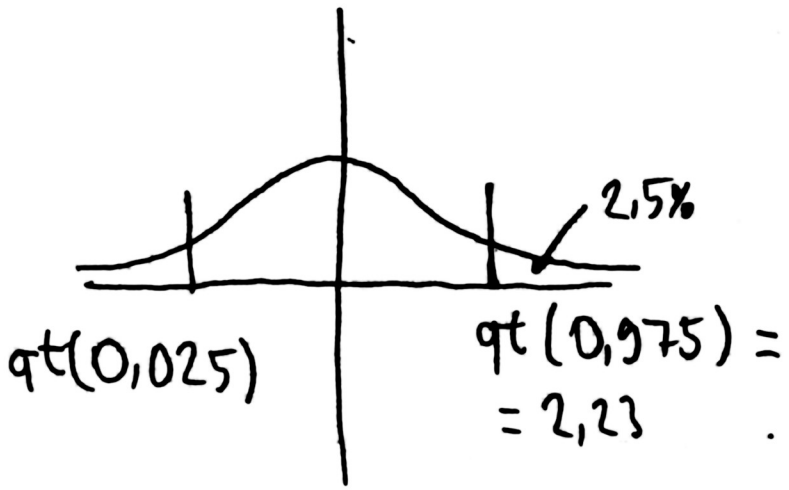
$A: EX \neq 0$

HODNOTA T TESTU:

$$T = \frac{\bar{X} - 0}{\sqrt{\frac{s^2}{n}}} = \underline{\underline{11,9}}$$

JE DALEKO VIAC AKO 2,23

$\Rightarrow$ HYPOTÉZU ZAMIETNEME



08.01.2008 Tu
PETER ČERNO

EX: MA1059 PRAVD. A STAT.

INTERVAL SPOLAHLIVOSTI

$\hat{EX} \pm q\, se(\bar{X}) \rightarrow$ STERODATNA CHYBA PRIEMERU

$\hat{EX} \pm q \sqrt{\frac{s^2}{n}}$

$q \leftarrow qt(0{,}975)$

$1320 \pm 2{,}23 \cdot \frac{366{,}7}{\sqrt{11}}$

$(1074,\ 1566)$

s 95% JE $EX = \hat{EX}$.

NEDOKÁZALI SME $EX \neq 1320$

PRE INTERVAL SPOLAHLIVOSTI $(\square, \infty)$

BY BOLO $\hat{EX} - q\sqrt{\frac{s^2}{n}}$

$q \leftarrow qt(0{,}95)$

$(1120, \infty)$

s 95% JE $EX = \hat{EX}$

NEDOKÁZALI SME $EX < 1320$

POČET SPADOV $\sim$ Poiss (λ).
ZA 100 DNÍ MÁME 2650 SPADOV.
MOŽNO TVRDIŤ, ŽE ZA DEŇ DOSTANIEM
a) V PRIEMERE VIAC AKO 25 SPADOV ?

MÁME $X_1, \dots, X_{100}$ ... POČTY SPADOV V JEDNOTLIVÝCH DŇOCH, $n = 100$

$X_i \sim$ Poiss (λ) SÚ NEZÁVISLÉ

$H: \; EX_i = 25$ $\qquad EX_i = \lambda$

$A: \; EX_i > 25$?

$$\cancel{\frac{\bar{X} - c}{\sqrt{\frac{S^2}{n}}} \sim N(0,1)}$$

NEPOZNÁME s

$$S_n \otimes = \sum_{i=1}^{100} X_i$$

PRI H (KEĎ JE $\lambda = 25$)

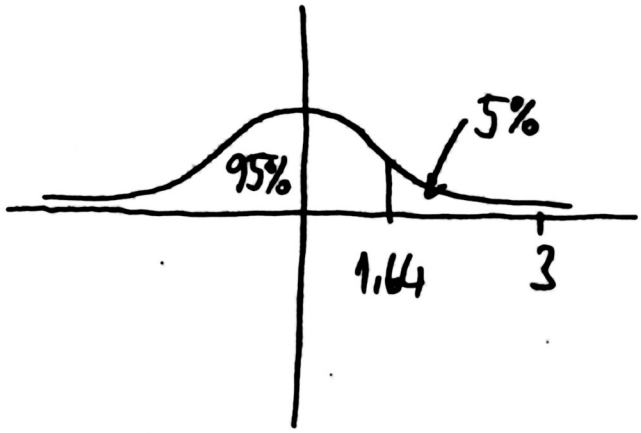
$$ES_n = \sum_{i=1}^{100} EX_i = 2500$$

$$\operatorname{var} S_n = \sum_{i=1}^{100} \operatorname{var} X_i = 100 \cdot \lambda = 2500$$

08.01.2008 Tu
PETER ČERNO

EX: MA1059 PRAVD. A ŠTAT.

$$Z = \frac{S_n - ES_n}{\sqrt{\operatorname{var} S_n}} = \frac{150}{50} = 3$$



$p = 0{,}0013$

ODPOVEĎ: PODARILO SA NÁM DOKÁZAŤ, ŽE DOSTÁVAME VIAC AKO 25 SPAMOV DENNE.

PRÍKLAD: REGRESIA

50 MLADÝCH MUŽOV SA UCHÁDZALO O PRÁCU U POLÍCIE. BOLO ZISŤOVANÉ % TUKU V TELE, HMOTNOSŤ A VÝŠKA.

$n = 50$, $Y = \%$ TUKU, $X =$ VÝŠKA

REGRESNÝ MODEL $Y = \beta_0 + \beta_1 X + \varepsilon$, $\varepsilon \sim N(0,1)$

		$\hat{se}(b)$
b_0	-55,91	24,96
b_1	0,39	0,14

RASTIE VÝZNAMNE % TUKU ? S VÝŠKOU

$H: \beta_1 = 0$

$A: \beta_1 > 0$

$$\frac{b-c}{se(b)} \sim N(0,1)$$

$$T = \frac{0,39}{0,14} = 2,79 > \underline{\underline{1,64}}$$

% TUKU RASTIE S VÝŠKOU

O KOĽKO MINIMÁLNE VZRASTIE % TUKU S 1 VÝŠKY ?

INTERVALOVÝ ODHAD

$$\boxed{b_1 \doteq q \; se(b_1)}$$

$$0,39 - 1,64 \cdot 0,14 \doteq 0,23$$

TJ. $(0,23, \infty)$

$\Rightarrow$ S 95% PRAVD. NA KAŽDÚ 1 VÝŠKY RASTIE % TUKU O 0,23.

08.01.2008 Tu
PETER ČERNO

EX: MA1059 PRAVD. A STATISTIKA

Y = % TUKU
X = HMOTNOST

$b_0 = -26{,}85$
$b_1 = 0{,}52$

Y = % TUKU
X_1 = VÝŠKA
X_2 = HMOTNOST

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

$b_0 = -13{,}29$
b_1 (VÝŠKA) = $\boxed{-0{,}27}$ PREČO ?
b_2 (HM) = $0{,}63$

07.01.2008 Mo
PETER ČERNO

MAI059 PRAVD. A ŠTATISTIKA

MOTTO: KTO UMÍ DĚLÁ, KTO NEUMÍ SIMULUJE!

- KEĎ NEUMÍME
- NIEČO SPOČÍTAME, CHCEME SI TO OVERIŤ
- SÚČASTNÉ POČÍTAČE A NUM. METÓDY NESTAČIA

- AKO GENEROVAŤ HODY MINCOU?

$$\begin{cases} 1 \,..\, p \\ 0 \,..\, 1-p \end{cases}$$

$$p_i = P(X = x_i)$$

VÝSTUPOM NAŠEHO ALGORITMU BY MALI BYŤ HODNOTY $\{x_i\}$

LEHNER (1951)

$$x_{n+1} = (a + b\, x_n) \bmod m$$