17. и 18. ПРИМЕНА СТАТИСТИКЕ У ПРОЦЕСУ КОНСТРУИСАЊА

РЕЗИМЕ:

Пошто се **статистички искази** ослањају на законе случаја и рачун вероватноће, важе само у оквиру извесне исказане поузданости. Код уобичајених техничких проблема рачуна се у општим случајевима са 95% исказане поузданости.

Уводни део:

Једна од особина свих чинилаца који одређују ефективност машинских система представља њихов случајни карактер и то се односи на све основне компоненте ефикасности: поузданост, готовост и погодност одржавања.

Због случајног карактера ефективности њено изучавање је засновано на методама теорије вероватноће и статистичким методама.

Ове **случајне величине**, без обзира на та да ли се покоравају дискретном (прекидном) или континуалном (непрекидном) закону промене, карактеришу одређени основни статистички показатељи. Под овим појмом се подразумевају вредности случајно промењљиве које показују спектар њених могућих промена, односно основне особине закона њене расподеле. Овакви показатељи су средња вредност, мода и медијална расподела, мере расипања око средње вредности, границе поверења и ранг појединих вредности.

Пошто се **статистички искази** ослањају на законе случаја и рачун вероватноће, важе само у оквиру извесне исказане поузданости. Код уобичајених техничких проблема рачуна се у општим случајевима са 95% исказане поузданости.

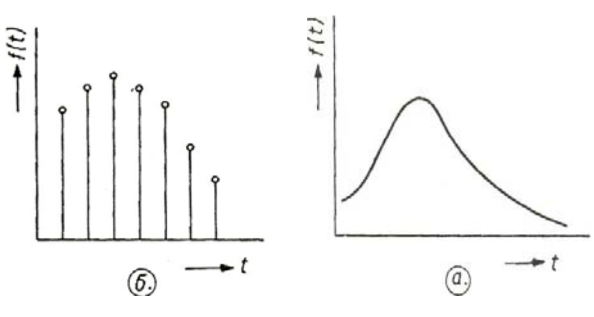
Средишњи део:

ОСНОВНЕ СТАТИСТИЧКЕ ВЕЛИЧИНЕ

Средња вредност популације (*m*), односно независно промењљиве *t* чија је густина расподеле **(СЛИКА)** дата функцијом *f(t)* одређена је изразом:

$$m = \int_{-\infty}^{+\infty} t \cdot f(t)dt$$

а. За непрекидне токове



$$m = \sum_{i=1}^{i=n} t_i\, p(t_i)$$

б. За прекидне токове

Где је:
t_i - средња вредност прекидне случајне величине
$p(t_i)$ – вероватноћа реализације величине t_i.

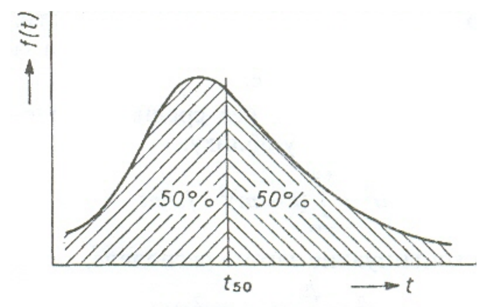
Ако се ради о ограниченом броју узорака као делу целе популације, средњу вредност посматране случајно промењљиве представља аритметичка средина:

$$m = \frac{\sum_{i=1}^{n} t_i}{n}$$ Где је n – укупан број података.

Овако одређена средња вредност је ближа средњој вредности целе популације, ако је број посматраних података већи.

Медијана

Медијана означава вредност независно промењљиве, чија је кумулативна вероватноћа реализације 0,5, што значи да је једнака вероватноћи да ће било који резултат бити мањи или већи од 0,5 (или 50%).За континуалне токове медијана или t_{50}, одређена је изразом: $0{,}5 = \int_{-\infty}^{t_{50}} f(t)dt$



МЕДИЈАНА

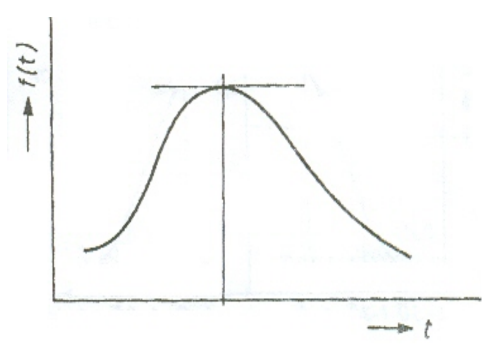
Значи, као горња граница интервала која одређује кумулативну вероватноћу од 0,5

Код узорака за које поседујемо само низ резултата, тј. за низ конкретних вредности посматране величине *t,* медијану представља вредност која се налази у средини свих резултата сложених по растућем реду.

Мода

Мода је вредност случајне величине која одговара највећој вероватноћи њене реализације, без обзира на то да ли се мења континуално или дискретно.

За континуално расподељену случајну величину t, чија је густина $f(t)$, вредност моде (модула) одређена је изразом: $\frac{df(t)}{dt} = 0$



МОДА

Као што се види из слике**, мода** одговара екстрему функције густине расподеле.

МЕРА РАСИПАЊА ОКО СРЕДЊЕ ВРЕДНОСТИ

За одређену популацију случајних величина поред средње вредности могуће је одредити и меру расипања независно промењљиве *t* око осе средње вредности. Овај показатељ је од посебног значаја за стохастичке процесе.

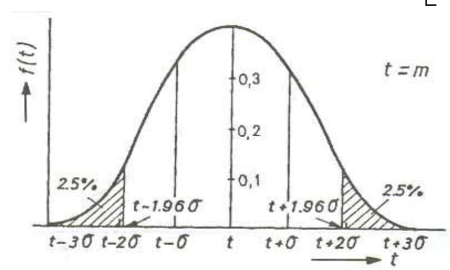
Мера расипања осе средње вредности изражава се у облику:

$$\sigma^2 = \int_{-\infty}^{+\infty} (t-m)^2 \cdot f(t)dt$$ За континуалне законе расподеле,

$$\sigma^2 = \sum_{i=1}^{n} (t_i - m)^2 \, p(t_i)$$ За дискретне случајно промењљиве расподеле.

Мера расипања се често изражава у облику тзв. Стандардне девијације $\sigma = +\sqrt{\sigma^2}$ а користи се за информације из једног ограниченог скупа података за који није познат закон расподеле.

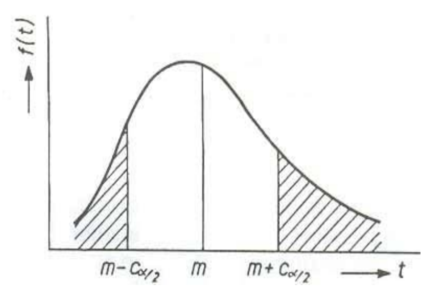
Процењена стандардна девијација (S) , тада се израчунава преко израза: $S = \left[\dfrac{\sum_{i=1}^{n} (t_i - m)^2}{n-1} \right]^{1/2}$



На слици је приказана **стандардна девијација нормалног закона расподеле**.

ГРАНИЦЕ ПОВЕРЕЊА

Одступање израчунате средње вредности од стварне средње вредности одговарајуће расподеле оцењује се на основу утврђивања граница, односно интервала поверења. Интервал поверења представља дијапазон у коме се са одређеном задатом вероватноћом налази стварна средња вредност, која одговара свим могућим реализацијама посматране случајне величине добијене кео резултат мерења. Исти смисао имају и границе поверења код утврђивања закона расподеле. Шематски приказ **граница и интервала поверења средње вредности** дат је на **слици .**



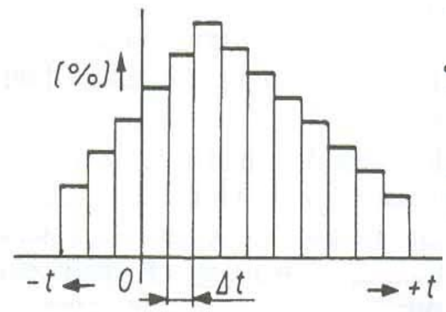
Са **слике** се види да границе поверења представљају вредности $C_{\alpha/2}$, које одговарају вероватноћи реализације у процентима, односно кумулативној вероватноћи α. Интервал поверења има облик:

$$m - C_{\alpha/2} \le m \le m + C_{\alpha/2}.$$

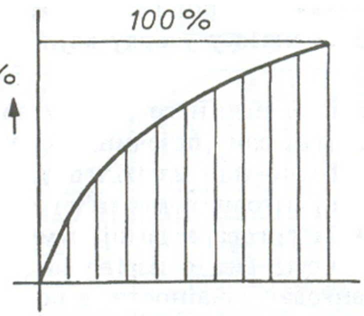
Вредности граница поверења могу се одредити на основу димензија узорака и величине стандардне девијације, односно одговарајуће процењене вредности. Величина $C_{\alpha/2}$ за одређену вероватноћу α, израчунава се из табеларних података за одговарајући закон расподеле (**види пример).**

ХИСТОГРАМ И ПОЛИГОН

Одређивањем основних статистичких величина добијају се информације о значајним показатељима посматраних узорака, али они ипак нису довољни за доношење закључака о популацији као целини. Зато је потребан виши ниво обраде статистичких величина које ће омогућити ближе упознавање стварног закона расподеле посматране случајно промењљиве величине. Прво приближавање овом циљу представља израчунавање релативних и кумулативних учесталости (фреквенције) и њихово гранично приказивање у облику **хистограма** и **полигона .**



ХИСТОГРАМ



ПОЛИГОН

За овакву обраду сви резултати мерења треба да се групишу у одређене класе, тј. интервале промене посматране величине, па се релативне учестаности одређују изразом

$$f_r = \frac{n_i}{n} \cdot 100\gamma\%$$

Где је:

n_i- број резултата мерења у свакој појединачној класи,

n – укупан број резултата мерења.

На основу релативних учестаности одређује се њихова кумулативна вредност f_q:

$$f_q = \sum_{i=1}^{n} \frac{n_i}{n} \cdot 100 y\%$$

Што се графички приказује у виду степенастог дијаграма или **полигона.**

На основу граничних приказа релативних и кумулативних учестаности може се или грубо или приближно просуђивати о карактеру закона расподеле посматране случајне величине. Осим тога, могу се оцењивати медијана и мода (расподеле), као и расипање око средње вредности.

Све ове оцене ипак нису довољна основа за објективно одлучивање. Зато је потребно ићи на више нивое обраде, односно тачније процењивање стварних закона расподеле.

Рангирање

Ако информације потичу од релативно малог броја узорака (мање од 50), хистограм је нетипичног карактера. Тада се користи метода рангирања, заснована на додељивању одговарајућег ранга сваком резултату понаособ, па се формирају кумулативне учестаности овако рангираних резултата.

Најзначајнија метода рангирања је одређивање тзв. Медијалног ранга (M_R) користећи емпиријски образац:

$$M_R = \frac{j - 0{,}3}{n + 0{,}4}$$

Где је: j – редни број резултата мерења,

n – укупан број резултата, односно број елемената у узорку који се посматра.

На исти начин могу се утврдити и други рангови појединих резултата. Најчешће се иде на рангове од 5% и 95%. Ти рангови означавају део популације која одговара датој кумулативној учестаности. Овако израчунате вредности рангова значајности, а посебно медијалног ранга, могу се приказати графички, аналогно степенастом дијаграму или полигону кумулативних фреквенција. Тако се добија знатно реалнија слика правог стања ствари, нарочито код релативно малог броја узорака.

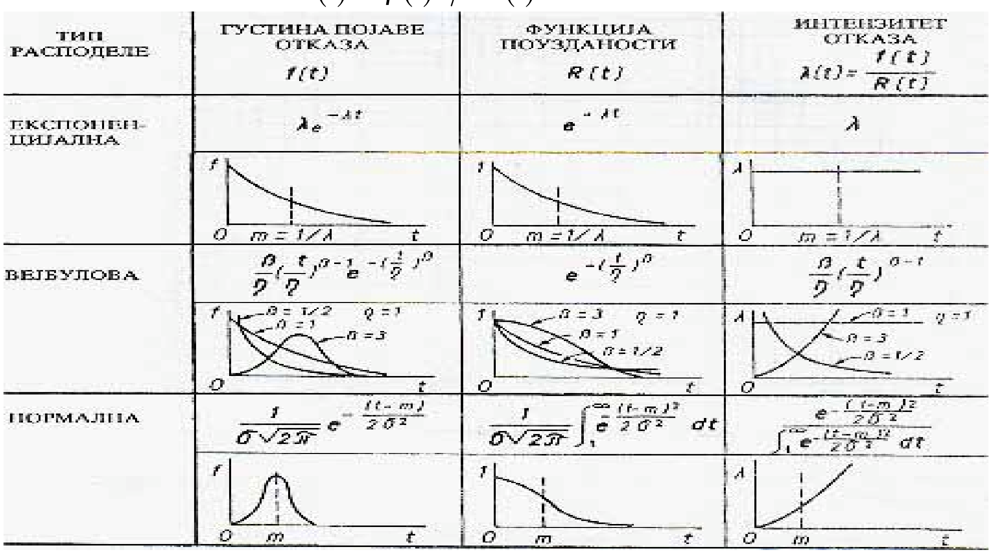
Теоријске расподеле вероватноће

Емпиријске расподеле поузданости могу се интерпретирати и теоријским расподелама вероватноће. По дефиницији, поузданост је једнака вероватноћи рада без отказа, односно:

$R(t) = \int_t^{\infty} f(t)dt$ Где је: *F(t)* – густина интервала времена рада до појаве отказа.

Непоузданост је: $F(t) = 1 - R\int_0^t f(t)dt$

А интензитет отказа: $\lambda(t) = f(t) \;/\; R(t)$



ТИП РАСПОДЕЛЕ	ГУСТИНА ПОЈАВЕ ОТКАЗА $f(t)$	ФУНКЦИЈА ПОУЗДАНОСТИ $R(t)$	ИНТЕНЗИТЕТ ОТКАЗА $\lambda(t) = \frac{f(t)}{R(t)}$
ЕКСПОНЕНЦИЈАЛНА	$\lambda e^{-\lambda t}$	$e^{-\lambda t}$	λ
ВЕЈБУЛОВА	$\frac{\beta}{\eta}(\frac{t}{\eta})^{\beta-1} e^{-(\frac{t}{\eta})^{\beta}}$	$e^{-(\frac{t}{\eta})^{\beta}}$	$\frac{\beta}{\eta}(\frac{t}{\eta})^{\beta-1}$
НОРМАЛНА	$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-m)}{2\sigma^2}}$	$\frac{1}{\sigma\sqrt{2\pi}}\int_t^{\infty} e^{-\frac{(t-m)^2}{2\sigma^2}} dt$	$\frac{e^{-\frac{(t-m)^2}{2\sigma^2}}}{\int_t^{\infty} e^{-\frac{(t-m)^2}{2\sigma^2}} dt}$

ТЕОРИЈСКА РАСПОДЕЛА ВЕРОВАТНОЋЕ

Највећу примену имају експоненцијална, **нормална и Вејбулова расподела** (слика**)** . Вејбулова расподела је параметарског типа, што јој даје велику флексибилност.

На слици су приказане основне карактеристике најчешће коришћених теоријских расподела, као и одговарајући аналитички изрази. У свим овим изразима t је случајно промењљива, односно време рада до појаве отказа.

Могућности интерпретације емпиријске расподеле поузданости неком теоријском расподелом најлакше могу да се провере помоћу вероватносних папира. Ти папири се конструишу за сваку теоријску расподелу. Уколико се подаци о кумулативним учестаностима отказа, унети у одређени вероватносни папир, налазе приближно на првој линији, нема разлога да се хипотеза о ваљаности тог теоријског закона не прихвати. Осим тога, подаци унети у вароватносни папир дају могућност да се одреде и сви параметри теоријског закона за тај случај.

Вејбулова расподела има два облика двопараметарски и тропараметарски.

Статистичка функција поузданости за двопараметарски облик је: $R(t) = e^{\left(\frac{t}{n}\right)^{\beta}}$

Где је:

β - параметар облика

μ- параметар размере

γ - параметар положаја.

Вредност параметара облика непосредно одређује облик функције густине. За β=1 Вејбулова двопараметарска расподела се у потпуности поклапа с експоненцијалним законом, односно:

$$R(t) = \exp\left(-\frac{t}{\mu}\right) = \exp(-\lambda t),$$

Што даје: $$\lambda = \frac{1}{m} = \frac{1}{\eta},$$

Где је:

η = m средња вредност расподеле.

За друге вредности параметра облика карактер Вејбулове расподеле се значајно мења. За β <1 расипање случајно промењљиве још је веће него код експоненцијалног закона, а за β ≈(2.5 до 3.5) Вејбулов закон се приближава нормалној расподели.

Параметар размере Вејбулове расподеле представља величину сразмерну средњој вредности расподеле $$m = \eta\Gamma\left(1 + \frac{1}{\beta}\right).$$

Вредности гама-функције $\Gamma\left(1 + \frac{1}{\beta}\right)$ за различите вредности параметара облика β дате су у прилогу.

Параметар положаја Вејбулове расподеле користи се за случајне промењљиве величине које не могу бити мање од неке минималне вредности. Овај параметар, у ствари, помера почетну тачку мерења на ту минималну вредност случајне промењљиве.

Вероватносни папир Вејбулове расподеле конструише се на следећи начин:

- двоструким логаритмовањем двопараметарског облика Вејбуловог закона добија се:

$$\ln\ln\frac{1}{l - F(t)} = \beta\ln t - \beta\eta$$

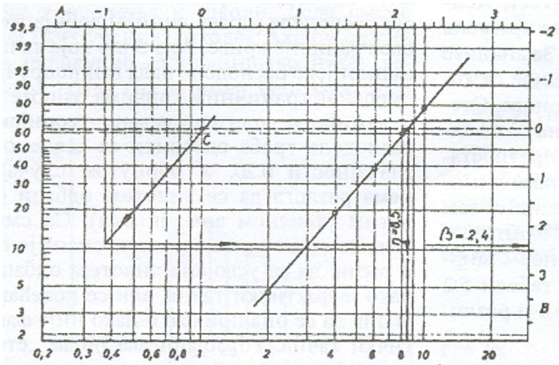
Увођењем замена

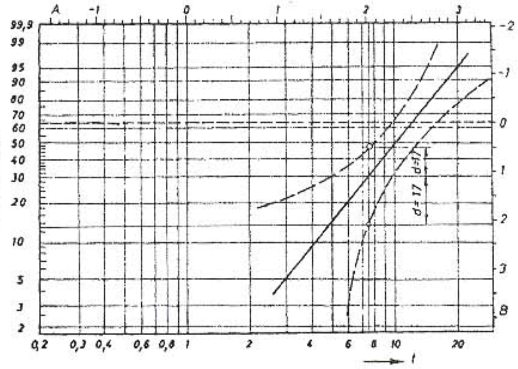
$$Y = \ln\ln\frac{1}{1 - F(t)}; x = \ln t;$$
$$a = \beta;\ b = -\beta\ln\eta$$

Добија се *Y=ax+b,* односно једначина праве линије.

Према томе, у координатном систему на чију је апсцису нанета случајна промењљива *t*, а на осу означену словом алгоритама случајно промењљиве, односно *x=lnt,* а на ординату величина *F(t),* односно на осу означену словом *B* величина *Y* , Вејбулова расподела је представљена правом линијом. Параметар облика представља нагиб ове праве (тангенте угла који права заклапа са *x* осом), док се параметар размера очитава у пресеку праве са линијом која одговара кумулативној учестаности случајне промењљиве од 63,2%.

Рад са вероватносним папиром Вејбулове расподеле објашњен је на примеру података датих у табели из које су кумулативне учестаности појаве отказа унете у вероватносни папир на слици**.**





ВЕЈБУЛОВА РАСПОДЕЛА – ВЕРОВАТНОСНИ ПАПИР **ТЕСТ КОЛМОГОНОВ - СМИРНОВЕ**

Табела :

Интервал мерне величине	Релативна учестаност %	Кумулативна учестаност %
0 – 2	4,5	4,5
2 – 4	14,5	19
4 – 6	20	39
6 – 8	21	60
8 – 10	18	78
10 – 12	8	86
Преко 12	14	100

Вид и се да све тачке одговарају правој линији. То даје основ за став да се поузданост посматраних елемената може интерпретираи овим законом расподеле. Значајност одступања појединих тачака од праве линије, што се практично увек догађа, треба проверити статистичким тестовима.

На примеру датом на слици за интерпретацију расподеле могу да се користе скале означене са **A** и **B.** Најпрактичније је да се из пресека обе нулте осе на скалама **A** и **B** (тачка **C**) повуче права паралелна усвојеној теоријској расподели, тј. правој која је повучена кроз унете тачке. Од пресека ове помоћне праве са вертикалом која одговара подељку (-1) на скали A, повлачи се хоризонтала, те се на скали B непосредно очитава величина параметара облика β, што је на слици назначено стрелицама.

Параметар размере непосредно одговара пресеку усвојене теоријске расподеле са хоризонталом која одговара кумулативној учестаности од 63,2%. Што се очитава на апсциси времена t . Са параметрима β и η, одређеним као на слици, функција поузданости за дати параметар може се изразити у облику:

$$R(t) = \exp\left[-\left(\frac{t}{8{,}5}\right)^{24}\right]$$

Док је на основу података из табеле за средњу вредност расподеле:

$$m = \Gamma\left(1+\frac{1}{\beta}\right)\eta = 8.887 \cdot 8{,}5 = 7.539$$

ТЕСТ КОЛМОГОНОВ - СМИРНОВЕ

Статистички тестови

Статистички тестови се користе за проверу да ли права повучена кроз тачке које одговарају експерименту, тј. емпиријској расподели у вероватносном папиру заиста одговара тој расподели.

	Степен значајности				
Бр.подат. ***n***	**0,20**	**0,15**	**0,10**	**0,05**	**0,01**

3	0,565	0,597	0,642	0,708	0,828
5	0,446	0,474	0,474	0,565	0,669
10	0,322	0,342	0,368	0,410	0,490
20	0,231	0,246	0,246	0,294	0,356
40	0,170	0,180	0,190	0,210	0,250
Преко 50	1,07/ $\sqrt{n}$	1,14/ $\sqrt{n}$	1,22/ $\sqrt{n}$	1,36/ $\sqrt{n}$	1,63/ $\sqrt{n}$

Постоји више оваквих тестова, као што су Колгомонов-Смирнове, *d* – тест, λ, Хенријева права и други, и њихова примена у статистици је шира. За анализе података о поузданости најчешће се користи тест Колгомонов-Смирнове. Степен сагласности се оцењује на бази одступања појединих тачака од претпостављене теоријске расподеле (праве линије), упоређивањем ових одступања са тзв. Критичним вредностима “d” . Критичне вредности “d” теста Колмогонов-Смирнове, дају се, како је дато у **табели** (извод), за различите податке n и различите степене значајности.

Пример примене **теста Колмогонов-Смирнове** дат је на **слици**, а односи се на скуп од 40 елемената, испитиваних до појаве отказа. За степен значајности 0.2 критична вредност “d” равна је 0.17. Наношењем ове вредности (у процентима) с обе стране претпостављене теоријске расподеле (праве линије) добијају се две граничне криве. Ако било која тачка емпиријске расподеле пада ван подручја омеђеног граничним кривима, хипотезу о важности претпостављене теоријске расподеле треба одбацити са степеном значајности 0.20. У обрнутом случају нема разлога да се хипотеза одбаци (с истим степеном значајности). Са смањенем степена значајности смањује се и ризик да се усвојена хипотеза одбаци иако је фактички тачна, али се повећава ризик да се она прихвати иако није фактички тачна. То објашњава појам "степена значајности" који се користи у овом тексту и другим статистичким тестовима.