STATISTIKA

Mjerenje, općenito uzevši, predstavlja promatranje u kojem se izvjesnom svojstvu objekta promatranja pridjeljuje rezultat, broj. Mjerenje se, dakako, ne mora odnositi samo na jedno svojstvo. U mnogim mjerenjima često dobivamo više rezultata istovremeno i u međusobnoj ovisnosti.

U svakom mjerenju očekuje se, a kod važnih mjerenja i zahtijeva, konzistentnost rezultata, tj. mogućnost ponavljanja dobivenog rezultata na istom ili istovjetnom objektu, promatrajući ista odnosno ekvivalentna svojstva. Ako ovaj uvjet nije ispunjen, rezultati imaju malu vrijednost. U zahtjevu za reproducibilnost rezultata susreće se, međutim, velik problem. S izuzetkom manjeg broja mjerenja (u nekim promatranjima s diskretnim varijablama) sva su mjerenja podložna varijacijama rezultata. Ove varijacije nazivaju se u našem jeziku pogreške, što je nesretno izabrano ime, jer je vezano za derogativnu riječ griješiti, kao da se nešto loše radi. Ako bismo bili neskromni, rekli bismo da se mjerenja katkada vrše loše. U francuskom i engleskom jeziku upotrebljava se bolja riječ koja je izvedena iz latinske riječi "erare" = lutati. Priroda svakog mjerenja je takva, da rezultati lutaju, rasipaju se, ako ne iz drugog onda iz razloga kojeg je već starogrčki filozof Heraklit jednostavno izrekao: "Ne možeš dva puta zagaziti u istu rijeku".

Problem konzistencije rezultata mjerenja postavlja se zbog toga u drugom vidu - koliko rezultati mjerenja mogu odstupati, a da ih još ne smatramo nekonzistentnim. Kriterije za konzistentnost rezultata daje nam matematička statistika i račun vjerojatnosti.

Nuklearni procesi, kao radioaktivni raspad, reakcije itd., podliježu statističkim zakonitostima kao i drugi procesi u prirodi. Nuklearne pretvorbe redovno su popraćene velikim oslobađanjem ili vezanjem energije, pa se mogu opažati pojedinačni događaji. To omogućuje proučavanje niza statističkih zakonitosti. U ovoj vježbi obradit će se najvažnije od tih zakonitosti.

POISSONOVA RASPODJELA

Poissonova raspodjela daje vjerojatnost u procesima u kojima imamo diskretnu varijablu, te malu i konstantnu vjerojatnost da će se pojedini događaj desiti. Može se primijeniti u vrlo različitim procesima, među kojima važno mjesto zauzimaju nuklearne pretvorbe.

Poissonova raspodjela može se izvesti iz Bernoullijeve raspodjele ako se načini granični prijelaz tako da broj mogućih događaja raste u beskonačnost, a prosječan broj događaja ostaje konstantan. Ovdje ćemo je izvesti na osnovi općih principa, koji vrijede za raspad radioaktivnih atoma:

1. Vjerojatnost za raspad u intervalu vremena određene duljine jednaka je za svaki atom dane vrste (atomi su identični).
2. Ta vjerojatnost za pojedini atom ne mijenja se ako se drugi atom raspadne (atomi su nazavisni).
3. Ta vjerojatnost neovisna je o vremenu (konstanta raspada ne ovisi o starosti atoma).
4. Ukupan broj atoma i broj vremenskih intervala promatranja je velik (vremenski intervali su kratki u usporedbi s prosječnim životom i statistička analiza je opravdana).

Pretpostavimo da je a prosječan broj događaja (npr. raspada) u jedinici vremena. U kratkom vremenskom intervalu između t i t+dt imat ćemo u prosjeku adt događaja. Ako je taj prosjek malen, tj. $adt \ll 1$, najčešće se neće opaziti ni jedan događaj, rijetko po jedan, dok je vjerojatnost za dva ili više događaja zanemariva. Budući da je

$$adt = 0 \cdot P(0,dt) + 1 \cdot P(1,dt) + 2 \cdot P(2,dt) + \ldots$$

uz zanemarivanje P(2,dt), P(3,dt) itd. imat ćemo

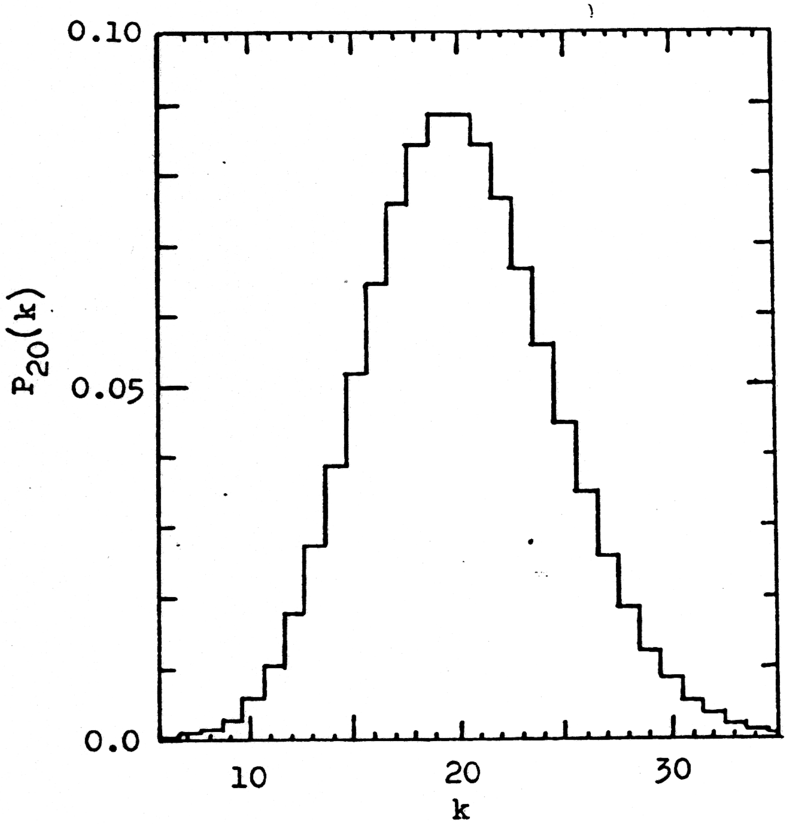
$$P(1,dt) = a\,dt, \qquad \text{i } P(0,dt) = 1 - a\,dt$$

Vjerojatnost da u vremenskom intervalu dužine t+dt opazimo k događaja jednaka je vjerojatnosti da u intervalu t opazimo k <u>i</u> u slijedećim dt nijedan, <u>ili</u> da u intervalu t opazimo k-1 i u slijedećih dt jedan događaj. Zbog kratkog intervala dt, kombinacije s dva ili više događaja u intervalu dt su zanemarene. Budući da "<u>i</u>" vjerojatnost znači produkt, a "<u>ili</u>" zbroj, slijedi

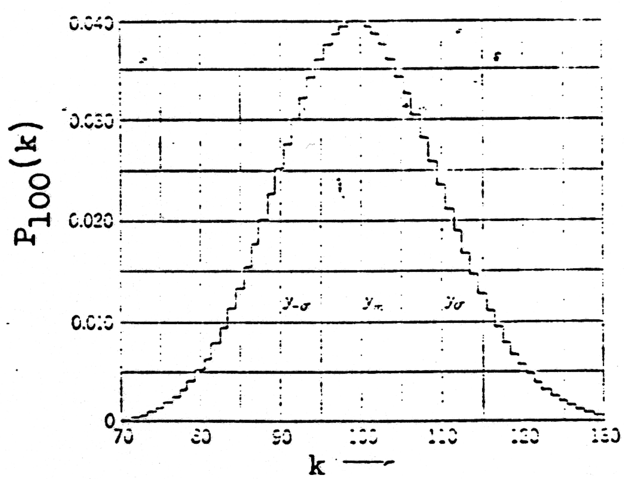
$$P(k,t + dt) = P(0,dt) \cdot P(k,t) + P(1,dt) \cdot P(k - 1,t)$$

odnosno

$$\frac{P(k,t + dt) - P(k,t)}{dt} = \frac{\partial P(k,t)}{\partial t} = a \left[P(k - 1,t) - P(k,t) \right]$$



Sl.1. Poissonova raspodjela za x = 20



Sl.2. Poissonova raspodjela za x = 100

Rješenje ove diferencijalne jednadžbe, uzevši u obzir da je k cjelobrojan, je Poissonova raspodjela

$$P_x(k) = \frac{x^k}{k!} e^{-x}$$

gdje je x = at prosječan broj dogadaja u intervalu duljine t. Poissonove raspodjele za x = 20 i x = 100 prikazuju Sl.1. i Sl.2. Za izračunavanje raspodjele praktično je primijeniti rekurzijsku formulu

$$P_x(0) = e^{-x}, \qquad P_x(k) = \frac{x}{k} P_x(k - 1)$$

Poissonova raspodjela je normirana sa srednjom vrijednošću x, varijancijom x i standardnom devijacijom $\sqrt{x}$:

$$\sum_{k=0}^{\infty} P_x(k) = e^{-x} \sum_k \frac{x^k}{k!} = 1$$

$$<k> = \sum_{k=0}^{\infty} k\, P_x(k) = e^{-x} \sum k \frac{x^k}{k!} = x$$

$$<(k - <k>)^2> = \sum k^2 \frac{x^k}{k!} e^{-x} - <k>^2 = x^2 + x - x^2 = x$$

Dakle, standardna devijacija iznosi

$$\sqrt{<(k - <k>)^2>} = \sqrt{x}$$

Kada se, dakle, u nekom eksperimentu dobije rezultat od k dogadaja, on predstavlja najbolju ocjenu srednje vrijednosti x, i ujedno najbolju ocjenu kvadrata standardne devijacije. Stoga se takav rezultat navodi

$$x = k \pm \sqrt{k}$$

s otprilike 68 % sigurnosti. Da je raspodjela normalna, interval sigurnosti bio bi 68.3 %.

Poissonova raspodjela ima svojstvo reproducibilnosti kod operacije preklopa pri zbrajanju, kompozicije. Za dvije diskretne raspodjele $R_1(k)$ i $R_2(l)$, gdje varijable raspodjela k i l primaju cjelobrojne vrijednosti $-\infty$, , -2, -1, 0, 1, 2, 3,..$+\infty$, operacija preklopa označuje novu raspodjelu u varijabli m koja je zbroj varijabli k i l. Npr., u jednom brojaču odbrojimo izvjestan broj impulsa k i u drugom l, od kojih bi svaki u ponavljanim mjerenjima bio raspodijeljen prema raspodjeli R_1, odnosno R_2. Pita se kakvu raspodjelu treba očekivati za sumu m = k+l. Ta je raspodjela dana sumom vjerojatnosti da prva varijabla ima jednu vrijednost k i varijabla l vrijednost m-k. Dakle,

$$R(m) = \sum_{-\infty}^{-\infty} R_1(k) \cdot R_2(m - k)$$

Budući da Poissonova raspodjela poprima vrijednost nula za negativne vrijednosti varijable, imat ćemo

$$R(m) = \sum_{0}^{m} P_x(k)\, P_y(m - k) = \sum_{0}^{m} \frac{x^k}{k!} e^{-x} \frac{y^{m-k}}{(m-k)!} e^{-y}$$

gdje su x i y srednje vrijednosti prve odnosno druge raspodjele. Gornji zbroj možemo pisati

$$R(m) = e^{-(x+y)} \frac{1}{m!} \sum_{0}^{m} \frac{m!}{k!(m-k)!} x^k y^{m-k} = \frac{(x+y)^m}{m!} e^{-(x+y)}$$

Dobiva se, dakle, ponovno Poissonova raspodjela, a njena srednja vrijednost jednaka je zbroju srednjih vrijednosti prve i druge raspodjele. Ovaj rezultat ujedno znači da dijeljenje jednog intervala mjerenja na više manjih intervala jednakog ukupnog trajanja neće u statističkom pogledu dati nove informacije o procesu. Naime, ako je jedno mjerenje s totalnim odbrojkom n i standardnom devijacijom $\sqrt{n}$ bilo razdijeljeno u nekoliko podrezultata n_1, n_2, ..., gdje je $n_1+n_2+n_3+\ldots = n$, pogreška svakog od podrezultata je $\sqrt{n_i}$. Standardna devijacija zbroja dana je s korijenom iz sume kvadrata pojedinih mjerenja, pa je u drugom slučaju rezultat,

$$\sum_i n_i \pm \sqrt{\sum_i (\sqrt{n_i})^2} = n \pm \sqrt{n}$$

dakle, jednak kao u neprekinutom mjerenju. Ovo ne znači da nema smisla dijeliti interval mjerenja na više manjih, jer su mogući drugi uzroci varijacije rezultata. Kod važnih mjerenja, radi provjere rada uređaja i konzistentnosti rezultata, uvijek se radi razdjela ukupnog vremena mjerenja na nekoliko manjih intervala. Ovi intervali pažljivo se raspoređuju za svako od posebnih mjerenja, npr. šum, promatranje jednog efekta, kalibracija, promatranje drugog efekta, šum itd.

Preklop dvije Poissonove raspodjele kod odbijanja rezultata (korelacija) ne daje ponovno Poissonovu raspodjelu. Ako je varijabla jedne raspodjele $k = 0, 1, 2, \ldots$, a druge $l = 0, 1, 2, \ldots$, raspodjela od $m = k-l$ bit će definirana od $-\infty$ do $+\infty$. U čestom slučaju odbijanja rezultata, koji su raspodijeljeni prema Poissonovoj raspodjeli, primjenjuju se pravila za normalnu raspodjelu. Poissonova raspodjela za veće srednje vrijednosti postaje približno normalna,

$$P_x(k) \rightarrow N(k) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(k-x)^2}{2\sigma^2}}$$

gdje je $\sigma = \sqrt{x}$.

Normalna raspodjela ima svojstvo da se u preklopu reproducira kod zbroja i kod razlike varijabla, a u oba slučaja varijancija (kvadrat standardne devijacije) se zbraja, $\sigma^2 = \sigma_1^2 + \sigma_2^2$. Tako se izračunavanje rezultata i standardne devijacije razlike dvaju odbrojaka, koji su raspodjeljeni prema Poissonovoj raspodjeli, može približno izračunati pomoću izraza

$$n_1 - n_2 \pm \sqrt{n_1 + n_2}$$

INTERVALNA RASPODJELA

Broj impulsa iz brojača u nekom intervalu vremena t, ako je efekt mrtvog vremena zanemariv, dan je Poissonovom raspodjelom. Srednja vrijednost jednaka je umnošku brzine brojanja a i duljine vremenskog intervala. Možemo se, međutim, zapitati kako će biti raspodijeljeno vrijeme između nekog početnog trenutka i trenutka kad stigne s-ti impuls u brojač. Brojila impulsa često se sastoje od brzog elektroničkog brojila i mnogo sporijeg mehaničkog brojila. Kod izvedbe tih brojila postavlja se pitanje, kako bliska u vremenu mogu biti dva odbrojka mehaničkog brojila nakon što se u elektroničkom brojilu brojenje reducira za faktor s. Taj faktor obično je 10^n, gdje je n broj dekada, ili 2^m, gdje je m broj binara u elektroničkom brojilu. Prosječni razmak povećava se za faktor s. Međutim, dobiva se još dodatno ujednačavanje na izlazu iz elektroničkog brojila, koje umanjuje ili potpuno uklanja gubitke brojanja mehaničkog brojila. Ovo ujednačavanje je posljedica promjene raspodjele vremenskih intervala kada se promatra razmak između s otkucaja umjesto raspodjele između dva susjedna otkucaja. Raspodjela vremena između s otkucaja naziva se opća intervalna raspodjela.

Intervalna raspodjela daje raspodjelu duljina vremenskih intervala između dva uzastopna događaja. Vjerojatnost da u vremenskom intervalu duljine t ,uz brzinu brojanja, a, ne dođe niti jedan otkucaj, jednaka je prema Poissonovoj raspodjeli:

$$P_{at}(0) = e^{-at},$$

budući da je za interval te duljine prosječan broja otkucaja at. Vjerojatnost da u slijedećem intervalu duljine dt stigne jedan otkucaj (pretpostavka je adt<<1), jednaka je adt. Prema tome, vjerojatnost da će duljina intervala bez otkucaja biti jednaka t jednaka je produktu ove dvije vjerojatnosti

$$dP_t = a\, e^{-at} dt$$

Ovo je intervalna raspodjela. Duljine vremenskih intervala raspodijeljene su prema eksponencijalnom zakonu. Lako se može (integracijom po vremenu od 0 do ∞) provjeriti, da je raspodjela normirana. Prosječna duljina vremenskog intervala između dva uzastopna otkucaja jednaka je

$$\langle t \rangle = \int_0^{\infty} t\, dP_t = \int_0^{\infty} t\, a\, e^{-a\,t}\, dt = 1/a,$$

dakle, recipročnoj vrijednosti brzine brojanja, kao što se moglo očekivati.

Opća s-terostruka intervalna raspodjela dobiva se ako se zahtijeva da u vremenskom intervalu t dođe s-1 impuls, a u slijedećem intervalu dt jedan. Ona je, dakle, produkt Poissonove raspodjele za k = s-1, uz prosječnu vrijednost x = at i Poissonove raspodjele za k = 1 uz prosječnu vrijednost a dt,

$$dP_t = \frac{(a\ t)^{s-1}}{(s-1)!}\, e^{-\,at}\, a\, dt$$

Integracijom po vremenu od 0 do ∞ može se provjeriti da je i opća intervalna raspodjela normirana. Prosječna vrijednost duljina s-terostrukog intervala jednaka je

$$\langle t\rangle = \int_0^\infty t\, dP_t = \int_0^\infty t\, \frac{(a\,t)^{s-1}}{(s-1)!}\, e^{-at}\, a\, dt = \frac{s}{a}$$

dakle, s puta prosječni razmak dvaju odbrojaka. Varijancija opće intervalne raspodjele jednaka je

$$\sigma^2 = \langle (t - \frac{s}{a})^2 \rangle = \int_0^\infty (t - \frac{s}{a})^2\, \frac{(a\,t)^{s-1}}{(s-1)!}\, e^{-at}\, a\, dt = \frac{s}{a^2}$$

Prema tome standardna devijacija jednaka je

$$\sigma = \frac{\sqrt{s}}{a} = \frac{\langle t\rangle}{\sqrt{s}}$$

Intervalnu raspodjelu za s = 1,2 i 4 prikazuje Sl.3., a prerađenu intervalnu raspodjelu s promjenom arugmenta u at/s, koja ima konstantnu srednju vrijednost, Sl.4.

Kada se rade mjerenja broja impulsa kroz interval vremena t, dobije se u prosjeku at otkucaja, a njihova standardna devijacija je $\sqrt{at}$. Kada ove otkucaje propustimo kroz elektroničko brojilo koje reducira broj za faktor s, standardna devijacija smanji se za taj faktor, ali još i za faktor $\sqrt{s}$, zbog efekta ujednačavanja koji je posljedica opće intervalne raspodjele. Kada promatramo uzastopne impulse, s = 1, tada za broj impulsa dobivamo standardnu devijaciju od Poissonove raspodjele.

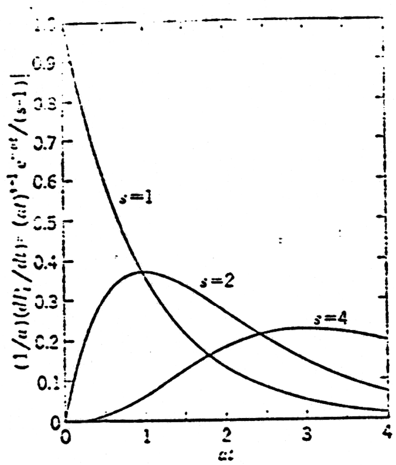
χ^2-KVADRAT RASPODJELA

Pearson je 1900. uveo χ^2-kvadrat (hi kvadrat) raspodjelu kao metodu za provjeru skladnosti između eksperimentalno mjerenih i teorijskih raspodjela. χ^2-kvadrat raspodjela je originalno izvedena na osnovi normalne raspodjele. Možemo je definirati kao

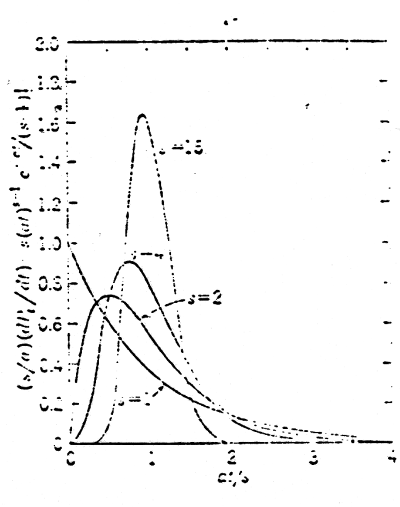
$$\chi^2 = \sum_i^N \frac{s_i^2}{\sigma_i^2}, \quad \text{broj stupnjeva slobode } F = N-1$$

ili

$$\chi^2 = \sum_i^N \frac{(x_i - x_{i,teor})^2}{\sigma_i^2},$$ broj stupnjeva slobode F = = N-(broj parametara teor. funkcije)



Sl.3. Opća s-terostruka intervalna raspodjela za s = 1, 2, i 4. Standardna devijacija raspodjele srazmjerna je $\sqrt{s}$.



Sl.4. Opća s-terostruka intervalna raspodjela za s = 1, 2, 4, i 16, prikazana u ovisnosti o renormaliziranoj varijabli at/s, kako bi se istakao efekt statističkog ujednačavanja. Širina ove raspodjele srazmjerna je $1/\sqrt{s}$.

gdje su x_i normalno raspodijeljeni rezultati mjerenja, s_i^2 najbolja ocjena varijancije svakog od tih rezultata, σ_i^2 varijancija, a $x_{i,teor}$ su teorijske (prilagodbene) vrijednosti. Iz same definicije vidimo da je u prosjeku

$$\langle\chi^2\rangle = F = \text{broj stupnjeva slobode}$$

U mjerenjima, koja su raspodijeljena prema Poissonovoj raspodjeli, postoji izravna relacija između prosječnog broja impulsa i varijancije: oni su jednaki. Ako je potrebno provjeriti raspodjelu, čija je varijabla raspodijeljena prema Poissonovoj raspodjeli, npr., broj intervala neke duljine u proučavanju intervalne raspodjele, u računanju χ^2 zamjenjuje se σ_i^2 s ocjenom koja se uzima jednaka teorijskoj vrijednosti broja događaja:

$$\chi^2 = \sum_1^n \left(\frac{(n_{exp} - n_{teor})^2}{n_{teor}}\right)_i , \quad F = n - 1$$

gdje je n_{exp} broj izmjerenih događaja (u provjeri intervalne raspodjele to je broj koliko puta se ponovio interval od t do $t+\Delta t$), a n_{teor} je broj koji se očekuje na osnovi teorijske raspodjele koja se provjerava.

Često se u nazivniku umjesto n_{teor} stavlja n_{exp}, s time da ako je $n_{exp} = 0$, onda se umjesto te vrijednosti uzima vrijednost 1. Ova zamjena uvelike olakšava traženje minimalne vrijednosti χ^2 podešavanjem parametera teorijske funkcije (one pomoću koje se računaju vrijednosti n_{teor}).

χ^2 raspodjele za manje vrijednosti broja stupnjeva slobode F dane su u tablicama i dijagramima. Kada je broj stupnjeva slobode velik, $F > 30$, χ^2 raspodjele mogu se aproksimirati s normalnom. Nalazi se da je $\sqrt{2\chi^2}$ u dobroj aproksimaciji normalna raspodjela sa srednjom vrijednošću jednakom 2F-1 i standardnom devijacijom jednakom 1.

Radni zadatak vježbe STATISTIKA

1. Ispitati pomoću slabog radioaktivnog izvora rad G.M. brojača i postaviti radni napon u sredinu platoa. Staviti brojač u štit. Podesiti brzinu brojanja na vrijednost koju će zadati voditelj praktikuma. Ako je potrebno, staviti slab izvor na položaj koji se neće mijenjati tijekom ovog niza mjerenja. Svakih 10 s bilježiti broj impulsa bez poništavanja brojila. Razlike daju broj impulsa u 10 s. Načiniti 100 očitanja. Nacrtati histogram: broj ponavljanja jednog broja impulsa u ovisnosti o broju impulsa. Izračunati teorijske frekvencije na osnovi Poissonove raspodjele i ucrtati u histogram. Ispitati slaganje eksperimentalnih i teoretskih rezultata pomoću χ^2 testa.

2. Podesiti brzinu brojanja na vrijednost koju će zadati voditelj praktikuma. Ako je potrebno, staviti slab izvor na položaj koji se neće mijenjati tijekom ovog niza mjerenja. Izvršiti 500 mjerenja duljine vremena između dva uzastopna impulsa pazeći da se kronometar uključuje s jednakim zakašnjenjima. Grupirati podatke u vremenske intervale 0-2,9 s, 3.0-5.9 s itd. Pritom ne zaustavljati brojilo, i istovremeno, pomoću sata , mjeriti vrijeme i tako odrediti brzinu brojanja. Nacrtati histogram: broj ponavljanja pojedinog vremenskog intervala u ovisnosti o rednom broju intervala. Na osnovi brzine brojanja i intervalne raspodjele izračunati teorijske frekvencije i unijeti ih u isti dijagram. Ispitati slaganje eksperimentalne i teorijske raspodjele pomoću χ^2 testa.

3. Postaviti uz G.M. brojač izvor koji daje oko 200 impulsa u sekundi. Izvršiti 100 mjerenja broja impulsa u jednoj sekundi, mjereći vrijeme kronometrom. Odrediti prosječan broj impulsa mjerenjem kroz 1000 sekundi. Izračunati χ^2 i ispitati da li je raspodjela dobivenih rezultata u skladu s Poissonovom raspodjelom. Raspraviti rezultat.