Distribusi Probabilitas Diskrit

Dadan Dasari

Daftar Isi

- DIstribusi Uniform
- Distribusi Binomial
- DIstribusi Multinomial
- Distribusi Hipergeometrik
- Distribusi Poisson

Distribusi Probabilitas Uniform Diskrit

Distribusi probabilitas yg paling sederhana adalah jikalau tiap nilai variabel random memiliki probabilitas yg sama untuk terpilih. Distribusi probabilitas spt ini diberi nama Distribusi Probabilitas Uniform Diskrit

Jika variabel random X bisa memiliki nilai $x_1, x_2, \ldots, x_k$ dan masing-masing bisa muncul dengan probabilitas yg sama maka distribusi probabilitasnya diberikan oleh :

$$f(x;k)=1/k \text{ untuk } x= x_1, x_2, \ldots, x_k$$

Notasi f(x;k) menyatakan nilai fungsi f tergantung pada k!

Contoh 1. Sebuah koin ideal memiliki muka : Angka dan Gambar. Jika x menyatakan banyaknya angka muncul, maka x=0,1 dan distribusi probabilitasnya

$$f(x;2)= \tfrac{1}{2} \qquad x=0,1$$

Distribusi Probabilitas Uniform Diskrit

Contoh 2.

Sebuah dadu ideal memiliki muka : 1,2,3,4,5,6. Jika x menyatakan mata dadu yg muncul, maka x= 1,2,3,4,5,6 dan distribusi probabilitasnya

$f(x;6)=1/6 \qquad x=1,2,3,4,5,6$

Contoh 3.

Sebuah kotak berisi 4 buah lampu masing-masing 40watt, 60watt, 100watt dan 500watt. Jika x menyatakan daya lampunya, dan diambil secara acak 1 lampu, maka distribusi probabilitasnya adalah:

$f(x;4)=1/4 \qquad x=40,60,100,500$

Mean dan Variansi Distribusi Uniform

Jika f(x;k) menyatakan distribusi probabilitas uniform , maka rata-ratanya

$$\mu = E[X] = \sum_{i=1}^{k} f(x_i;k)x_i = \frac{1}{k}\sum_{i=1}^{k} x_i$$

dan variansinya :

$$\sigma^2 = E[(X-\mu)^2] = \sum_{i=1}^{k} f(x_i;k)(x_i-\mu)^2 = \frac{1}{k}\sum_{i=1}^{k}(x_i-\mu)^2$$

Contoh.

Hitunglah nilai rata-rata mata dadu yg keluar dan variansinya!

$$\mu = \frac{1}{6}\sum_{i=1}^{6} x_i = \frac{1+2+3+4+5+6}{6} = 3.5$$

$$\sigma^2 = \frac{1}{6}\sum_{i=1}^{6}(x_i-\mu)^2 =$$

$$= \frac{(1-3.5)^2+(2-3.5)^2+(3-3.5)^2+(4-3.5)^2+(5-3.5)^2+(6-3.5)^2}{6} = 2.92$$

Distribusi Binomial: Proses Bernoulli

Proses Bernoulli adalah, sebuah proses eksperimen statistik yg memiliki ciri-ciri:

- Eksperimen terdiri dari n kali pengulangan
- Tiap kali, outcome hanya dua macam, dilabeli “sukses” dan “gagal”
- Probabilitas “sukses” di tiap percobaan, p, besarnya tetap dari satu percobaan ke berikutnya.
- Satu percobaan dan yg berikutnya bersifat independen

Contoh.

Sebuah proses Bernoulli untuk QC dilakukan dengan memilih 3 komponen secara simultan dari sebuah proses produksi. Setiap komponen yg diambil dinyatakan “sukses” jika ternyata rusak, dan “gagal” jika ternyata komponen tsb baik (sebenarnya boleh juga definisinya dibalik!). Variabel random X didefinisikan sebagai banyaknya “sukses” dalam pengambilan 3 komponen tsb.

Distribusi Binomial: Proses Bernoulli

Contoh (lanjutan).

Ruang sampel bagi X adalah (S: sukses, G:gagal):

Outcome	SSS	SSG	SGS	SGG	GSS	GSG	GGS	GGG
X	3	2	2	1	2	1	1	0

Misalkan diketahui dari masa lalu, sebanyak 25% produksi komponen tersebut rusak ("S"). Jadi probabilitas 1 kali pengambilan menghasilkan rusak = probabilitas "sukses" = p= ¼, berarti probabilitas "gagal" = 1- ¼ = ¾ .

Sebagai contoh probabilitas outcome= SSG

p(SSG)=p(S)p(S)p(G)= ¼* ¼ * ¾ = 3/64, jadi untuk X=2, ada 3 outcome yg terkait : SSG, SGS, GSS, maka jika f(X=2) menyatakan probabilitas X=2,

f (X=2) = 3*3/64 = 9/64.

Distribusi Binomial: Proses Bernoulli

Contoh (lanjutan).

Dengan cara yg sama bisa diturunkan probabilitas untuk X=0,1 dan 3, dan hasilnya adalah fungsi distribusi probabilitas f(x) sbb:

X	0	1	2	3
f(X)	27/64	27/64	9/64	1/64

Variabel random X ini disebut variabel random binomial, sedangkan fungsi distribusinya f(x) disebut fungsi distribusi binomial, dan dituliskan sbb:

$$f(x) = b(x;n,p)$$

Untuk menegaskan bahwa probabilitas x ditentukan oleh banyak eksperimennya (n, dalam contoh di atas n=3), dan bergantung pada probabilitas sukses di tiap eksperimen (p).

Jadi f(x=2) =b(2;3,0.25) = 9/64

Distribusi Binomial: General Case

Kasus distribusi binomial umum:

- dilakukan eksperimen sebanyak n kali pengambilan
- dari n tsb, sebanyak x dikategorikan “sukses”, jadi sebanyak n-x adalah “gagal”
- probabilitas “sukses” di tiap percobaan = p, berarti probabilitas “gagal “, q=1-p.

Maka probabilitas terjadinya outcome dengan konfigurasi x “sukses” dan (n-x) “gagal” **tertentu**, adalah:

$P(SSS \ldots GGG) = ppp \ldots qqq = p^x q^{n-x}$

Sebab S ada x buah dan G sebanyak (n-x) buah. Tentu ada banyak konfigurasi lain yg juga memiliki x buah S dan (n-x) buah G. Sehingga probabilitas mendapatkan hasil eksperimen yg memiliki x buah S dan (n-x) buah G adalah:

$$C^n_x \, p^x q^{n-x} \quad = b(x;n,p)$$

Fungsi Distribusi Binomial

Proses Bernoulli dimana tiap eksperimen (pengambilan) memiliki probabilitas sukses p (atau probabilitas gagal q=1-p). Maka fungsi distribusi probabilitas f(x) yang menyatakan dari n kali eksperimen (pengambilan) yg independen mengandung x buah yg sukses adalah :

$$b(x; n, p) = \binom{n}{x} p^x q^{n-x}$$

Dengan x=0,1,2,….,n

Contoh.

Probabilitas sebuah komponen mobil tidak rusak ketika dijatuhkan adalah ¾. Berapakah probabilitasnya ada 2 dari 4 komponen yg dijatuhkan akan tidak rusak.

Contoh.

Jawab.

Misal kita definisikan “sukses” = tidak rusak, probabilitas “sukses”, p=3/4. Jadi probabilitas “gagal, q= 1-3/4 = ¼. Total percobaan ada n=4, jumlah yg tidak rusak, “sukses”, x=2. Jadi probabilitas 2 dari 4 komponen yg dijatuhkan tidak rusak diberikan oleh:

$$b(x=2;n=4,p=\frac{3}{4})=\binom{n}{x}p^x q^{n-x}=\binom{4}{2}(\frac{3}{4})^2(\frac{1}{4})^{4-2}=\frac{4!}{2!(4-2)!}\frac{9}{16}\frac{1}{16}=\frac{27}{128}$$

Sifat dari b(x;n,p) sebagai fungsi distribusi probabilitas adalah:

$$\sum_{x=0}^{n} b(x;n,p)=1$$

Karena seringkali kita memerlukan probabilitas untuk X dalam sebuah interval, misal P(X<r) atau P(a<X≤b) maka, dibuat tabel fungsi distribusi binomial kumulatif sbb:

$$B(r;n,p)=\sum_{x=0}^{r} b(x;n,p)$$

Tabel Distribusi Probabilitas Binomial Kumulatif

Table A.1 Binomial Probability Sums $\sum_{x=0}^{r} b(x;n,p)$

		p						
n	*r*	0.10	0.20	0.25	0.30	0.40	0.50	0.60
1	0	0.9000	0.8000	0.7500	0.7000	0.6000	0.5000	0.4000
	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	0	0.8100	0.6400	0.5625	0.4900	0.3600	0.2500	0.1600
	1	0.9900	0.9600	0.9375	[illegible]	0.8400	0.7500	0.6400
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	0	0.7290	0.5120	0.4219	0.3430	0.2160	0.1250	0.0640
	1	0.9720	0.8960	0.8438	0.7840	0.6480	0.5000	0.3520
	2	0.9990	0.9920	0.9844	0.9730	0.9360	0.8750	0.7840
	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

$$B(r;n,p)=\sum_{x=0}^{r} b(x;n,p)$$

B(r=1;n=2,p=0.30) = 0.9100

Contoh.

Probabilitas seorang pasien yg sakit suatu penyakit flu sembuh adalah 40%. Jikalau 15 orang diketahui telah tertular penyakit ini, berapakah probabilitasnya bahwa (a) paling tidak 10 orang sembuh, (b) antara 3 hingga 8 orang sembuh (c)tepat 5 orang sembuh?

Jawab.

Ini adalah proses Bernoulli. Probabilitas “sukses”, yaitu sembuh adalah $p = 0.4$. Variabel random X menyatakan banyak orang yang “sukses” = sembuh, sedangkan total percobaannya adalah $n=15$.

a) P (paling tidak 10 sembuh) = $P(X \geq 10) = 1 - P(X<10) =$

$= 1 - B(r=9; n=15, p=0.4) = 1 – 0.9662 = 0.0338$

b) P (antara 3 sd 8 sembuh) = $P(3 \leq X \leq 8) = P(X \leq 8) - P(X<3) =$

$= B(r=8; n=15, p=0.4) - B(r=2; n=15, p=0.4) = 0.9050 - 0.0271 = 0.8779$

c) P (5 sembuh) = $P(X=5) = P(X \leq 5) - P(X<5) =$

$= B(r=5; n=15, p=0.4) - B(r=4; n=15, p=0.4) = 0.4032 - 0.2173 = 0.1859$

Mean dan Variansi Distribusi Binomial

Jika X adalah variabel dg distribusi binomial b(x;n,p), maka mean dan variansinya adalah:

$$\mu = np \quad \text{dan } \sigma^2 = npq$$

Bukti:

Misalkan kita lakukan percobaan sebanyak n kali. Tiap kali outcomenya disebut I_k yg bisa bernilai "sukses" atau "gagal" dengan probabilitas "sukses" = p. Maka variabel random X yg menyatakan jumlah "sukses" dari n eksperimen akan memiliki mean:

$\mu = E(X) = E\,(I_1 + I_2 + I_3 + \ldots. + I_n) = E(I_1) + E(I_2) + \ldots + E(I_n) =$

$\mu = p + p + ,,, + p = np$

Contoh.

Probabilitas seorang pasien yg sakit suatu penyakit flu sembuh adalah 40%. Jikalau 15 orang diketahui telah tertular penyakit ini, (a) Berapakah rata-rata jumlah orang yg sembuh? (b) Menurut teorem Chebysev paling tidak sebanyak 75% kasus akan jatuh dalam interval $\mu - 2\sigma < X < \mu + 2\sigma$. Terapkan dalam kasus ini dan beri interpretasi.

Jawab.

a) Dalam kasus ini probabilitas sembuh, p=0.4, banyak percobaan, n=15, sehingga rata-rata jumlah orang yang sembuh

$$\mu = np = 15*0.4 = 6 \text{ orang}$$

b) Variansinya :

$\sigma^2 = npq = np(1-p) = 15*(0.4)(1-0.4) = 3.6$ dengan STD $\sigma = 1.897$,

$\mu - 2\sigma = 6 - 2(1{,}897) = 2.206$ dan $\mu + 2\sigma = 6 + 2(1{,}897) = 9.794$.

Artinya (menurut Chebysev) terdapat probabilitas paling tidak 75% pasien yg sembuh jumlahnya antara 2.206 s/d 9.794 atau dibulatkan antara 3 s.d 9.

Contoh.

Diperkirakan 30% sumur di sebuah desa tercemar.Untuk memeriksa kebenaran hal tsb dilakukan pemeriksaan dengan secara acak mengambil 10 sumur.

(a) Jika perkiraan tsb benar, berapakah probabilitasnya tepat 3 sumur tercemar?

(b) Pertanyaan yg sama tapi lebih dari 3 sumur yg tercemar?

Jawab.

Probabilitas 1 sumur tercemar p=0.3 ("sukses"), jadi probabilitas tidak tercemar ("gagal") q=1-p = 1-0.3=0.7.

Total pengambilan n=10 buah.

a) Tepat 3 sumur tercemar, x=3.

$P(x=3;n=10,p=0.3)= B(r=3;n=10,p=0.3)-B(r=2;n=10,p=0.3)$

$= 0.6496 – 0.3838 = 0.2668$ (27%).

b) Lebih dari 3 sumur tercemar x>3,

$P(x>3;n=10,p=0.3)= 1- P(x\leq 3;n=10,p=0.3)=$

$= 1 – B(r=3;n=10,p=0.3) =1 – 0.6496 = 0.3504 = 35\%$

Soal.

Soal yg sama. Misalkan ternyata dari 10 sampel yg diambil secara acak sebanyak 6 mengandung pencemaran. Pergunakanlah perhitungan probabilitas, untuk memberik komentar ttg kemungkinan hal spt terjadi, jikalau perkiraan semula benar.

Distribusi Multinomial

Sebagai generalisasi dari distribusi binomial adalah dengan melonggarkan kriteria banyaknya outcome yg mungkin jadi > 2. Dalam hal ini maka percobaannya disebut percobaan multinomial sedangkan distribusi probabilitasnya disebut distribusi multinomial.

Definisi:

Misal setiap percobaan bisa menghasilkan k outcome yg berbeda, E_1, E_2, …,E_k masing-,masing dengan probabiliitas p_1, p_2, …,p_k. Maka distribusi multinomial $f(x_1,x_2,\ldots,x_k; p_1,p_2, \ldots,p_k, n)$ akan memberikan probabilitas bahwa E_1 akan muncul sebanyak x_1 kali, E_2 akan muncul sebanyak x_2 kali, dst dalam pengambilan independen sebanyak n kali, jadi

$$x_1+ x_2+ \ldots + x_k=n$$

dengan $p_1+p_2+ \ldots + p_k =1$ dan

$$f(x_1,\ldots,x_k; p_1,\ldots,p_k,n)=\binom{n}{x_1,x_2,\ldots x_k} p_1^{x_1} p_2^{x_2}\ldots p_k^{x_k}$$

Contoh.

Sebuah airport memiliki 3 buah landas pacu (runway), dan probabilitas sebuah runway dipilih oleh pesawat yg akan mendarat adalah:

runway -1 : 2/9

runway -2 : 1/6

runway -3 : 11/18

Berapakah probabilitas 6 pesawat yg datang secara acak di distribusikan ke dalam runway-runway tsb spt berikut:

runway -1 : 2 pesawat

runway -2 : 1 pesawat

runway -3 : 3 pesawat

Jawab.

Pemilihan runway acak dan independen, dengan p_1=2/9, p_2=1/6 dan p_3=11/18. Probabilitas untuk x_1=2, x_2= 1 dan x_3=3 adalah

$$f(x_1 = 2, x_1 = 1, x_3 = 3; p_1 = \frac{2}{9}, p_1 = \frac{1}{6}\ p_1 = \frac{11}{18}, n = 6) = \binom{6}{2,1,3} (\frac{2}{9})^2 (\frac{1}{6})^1 (\frac{11}{18})^3 = 0.1127$$

Distribusi Hipergeometrik

Distribusi Hipergeometrik sangat serupa dengan distribusi binomial,

Persamaannya:

Keduanya menyatakan probabilitas sejumlah tertentu percobaan masuk dalam kategori tertentu.

Perbedaannya:

- Binomial mengharuskan ketidakbergantungan dari satu percobaan (trial) ke percobaan berikutnya.
- Jadi sampling harus dilakukan dengan dikembalikan (replaced)
- Hipergeometrik tidak mengharuskan ketidakbergantungan, jadi sampling dilakukan tanpa mengembalikan outcome yg sudah keluar.

Distribusi Hipergeometrik

Distribusi Hipergeometrik sangat serupa dengan distribusi binomial,

Persamaannya:

Keduanya menyatakan probabilitas sejumlah tertentu percobaan masuk dalam kategori tertentu.

Perbedaannya:

- Binomial mengharuskan ketidakbergantungan dari satu percobaan (trial) ke percobaan berikutnya.
- Jadi sampling harus dilakukan dengan dikembalikan (replaced)
- Hipergeometrik tidak mengharuskan ketidakbergantungan, jadi sampling dilakukan tanpa mengembalikan outcome yg sudah keluar.

Distribusi Hipergeometrik: Definisi

Distribusi Hipergeometrik dari variabel random X yang menyatakan banyaknya outcome yang "sukses" dari sampel random sebanyak n yg diambil dari populasi sebanyak N, dimana dari N tsb sebanyak k buah adalah "sukses" dan sisanya "N-k" adalah "gagal":

$$h(x;N,n,k)=\frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

Suku pembagi (denominator) menyatakan banyak kombinasi yg terjadi jika dari N obyek diambil n tiap kali.

Faktor pertama suku terbagi (numerator) menyatakan banyaknya kombinasi dari obyek berjenis "sukses" yg berjumlah k jika tiap kali diambil sebanyak x buah.

Faktor kedua suku terbagi (numerator) menyatakan banyaknya kombinasi dari obyek berjenis "gagal" sebanyak N-k jika tiap kali diambil sebanyak (n-x) buah.

Contoh DIstribusi Hipergeometrik Dalam Sampling Plan.

Suatu jenis suku cadang mobil dijual dalam bentuk paket yg isinya 10 buah. Produsen merasa bahwa bahwa paket tsb dinyatakan “dapat diterima” jikalau tak lebih dari 1 buah suku cadang/paket yg cacat. Untuk memeriksa kualitasnya dilakukan sampling secara random diambil beberapa paket, dan ditiap paket dilakukan pemeriksaan terhadap 3 buah suku cadang dari paket yg disampel. Kemudian paket dinyatakan baik jika dari 3 yg diperiksa tsb tidak satupun yg cacat. Berapakah probabilitasnya seandainya sampel yg diambil sebenarnya mengandung 2 buah suku cadang cacat (jadi unacceptable), tapi ketika diambil sampel 3 ternyata tak satupun juga cacat, sehingga salah mengambil kesimpulan!

Jawab.

Misalkan bahwa ada lot yg benar-benar tak bisa diterima, karena 2 dari 10 isinya cacat. Kita hitung berapa probabilitasnya bahwa teknik sampling yg kita lakukan dapat menemukan hal ini.

Contoh DIstribusi Hipergeometrik Dalam Sampling Plan.

Misal X adalah banyak suku cadang yg cacat, maka probabilitas bahwa dari 3 suku cadang yg diambil tak satupun cacat adalah sbb:

Jumlah yg cacat di paket k=2, yg terambil tidak ada, X=0. Isi satu paket N=10, jadi yg baik N-k=10-2=8. Dari paket diambil n=3 sampel.

Banyaknya kombinasi bahwa dari k=2 cacat di paket tidak terambil sama sekali (x=0) adalah C^2_0 = 2!/(0!2!)=1. Dan kombinasi dari 8 yg cacat diambil 3 buah ada sebanyak C^8_3 = 8!/(3!5!) = 8x7x6/6=56.

Sedangkan kalau dari 10 diambil 3 buah item, banyak kombinasi iterm yg mungkin adalah C^{10}_3 = 10!/7!3! = 10x9x8/6 = 120.

Jadi probabilitas bahwa yg terambil mengandung 3 buah item dan tak satupun cacat adalah :

$$h(x=0; N=10, n=3, k=2) = \frac{\binom{2}{0}\binom{8}{3}}{\binom{10}{3}} = \frac{1*56}{120} = 0.467 = 47\%$$

Contoh DIstribusi Hipergeometrik Dalam Sampling Plan.

Jadi ada probabilitas 47% bahwa walaupun sebenarnya paketnya mengandung 2 cacat, tapi dari 3 sampel suku cadang yg diperiksa tak satupun juga yg cacat, sehingga salah mengambil kesimpulan bahwa paket tsb bagus! (Acceptable). Berarti sampling plan seperti ini JELEK! Sangat mungkin mengambil kesimpulan yg salah.

DIstribusi Hipergeometrik: Mean dan Variansi

Jika h(x;N,n,k) menyatakan distribusi hipergeometrik untuk variabel acak x, yg menyatakan jumlah item yg “sukses” bilamana dari N item diambil sebanyak n buak, dan sebenarnya sebanyak k item sukses dari N buah tsb, maka rata-rata x diberikan oleh:

$$\mu = \frac{nk}{N}$$

Dan variansinya oleh:

$$\sigma^2 = \frac{N-n}{N-1} \bullet n \bullet \frac{k}{N}(1 - \frac{k}{N})$$

DIstribusi Hipergeometrik: Mean dan Variansi

Contoh.

Paket yg terdiri dari 40 item, dinyatakan ditolak jikalau paket tsb mengandung item cacat 3 atau lebih. Prosedur sampling yg diterapkan adalah dengan mengambil sampel 5 item, dan memeriksanya jikalau ditemui yg cacat, maka keseluruhan paket ditolak. (a) berapakah probabilitasnya bahwa jika ternyata paket mengandung 3 item cacat, bisa ketemu 1 cacat di sampel 5 item yg dipilih? (b) jika X menyatakan banyak item yg cacat, hitunglah mean dan variansi, (c) pergunakan teorema Chebysev untuk menaksir interval $\mu \pm 2\sigma$.

DIstribusi Hipergeometrik: Mean dan Variansi

Jawab:

a) Banyak item cacat terambil x=1, banyak total item N=40, sampel yg diambil n=5, total item cacat di populasi k=3. Probabilitas 1 item cacat terambil dari 5 yg diambil:

h(x=1;N=40,n=5,k=3) =

$$h(x=1; N=40, n=5, k=2) = \frac{\binom{3}{1}\binom{37}{4}}{\binom{40}{5}} = 0.3011 = 30\%$$

Jadi sampling plan ini tak baik, sebab hanya mampu mendeteksi cacat dengan probabilitas 30%.

b) Nilai mean x yaitu rata-rata jumlah sampel cacat yg terambil dan variansinya adalah:

$$\mu = k\frac{n}{N} = 3*\frac{5}{40} = 0.375 \qquad \sigma^2 = \frac{N-n}{N-1}n \bullet \frac{k}{N}(1-\frac{k}{N}) = \frac{40-5}{40-1}5 \bullet \frac{3}{40}(1-\frac{3}{40}) = 0.3113$$

DIstribusi Hipergeometrik: Mean dan Variansi

Jawab:

c) Standard deviasinya $\sigma = 0.558$, sehingga interval $\mu \pm 2\sigma$ adalah:

$0.375 \pm 2(0.558) = -0.741$ s/d 1.491. Teorema Chebysev menyatakan terdapat probabilitas 75% dari sampel 5 yg diambil tersebut akan mengandung jumlah item yg cacat sebanyak -0.741 dan 1.491. Jadi berarti 3 dari 4 kesempatan, dari 5 buah sampel item yg diambil mengandung komponen yg cacat kurang dari 2.

DIstribusi Hipergeometrik vs Binomial

Jikalau ukuran sampel diambil n jauh lebih kecil dari ukuran populasinya N (n«N) maka distribusi hipergeometrik akan sangat mirip dengan distribusi binomial dengan:

k/N memainkan peranan probabilitas “sukses” binomial p, sehingga mean dan variansinya mengikuti distribusi binomial yaitu:

$$\mu = k\frac{n}{N} = n * p \qquad \sigma^2 = 1 * n * p * q = n\frac{k}{N}(1 - \frac{k}{N})$$

Sebagai pedoman praktis, seringkali dipergunakan jikalau n/N < 5% maka dipergunakan distribusi binomial sebagai pengganti distribusi hipergeometrik.

DIstribusi Hipergeometrik vs Binomial

Contoh.

Sebuah pabrik ban menyatakan dari 5000 ban yang dikirim ke distributor sebanyak 1000 warnanya sedikit pudar. Seorang pelanggan membeli 10 ban dari distributor secara acak saja. Berapa probabilitasnya bahwa ada 3 buah ban yg warnanya sedikit pudar?

Jawab: Populasinya N=5000, ukuran sampelnya n=10 ($n/N < 5\%$), jadi bisa dipakai distribusi binomial saja, dengan probabilitas warna sedikit pudar $p=k/N = 1000/5000 = 0.2$, dan tidak pudar $q=1-p=0.8$. Jumlah sampel n=10, banyak yg pudar x=3, berarti probabilitasnya :

$$P(x=3;n=10,p=0.2)= B(r\leq 3;n=10,p=0.2)-B(r\leq 2;n=10,p=0.2)$$

$$= 0.8791 -0.6778 = 0.2013 = 20\%$$

Periksalah, jika dipergunkan distribusi hipergeometrik hasilnya=0.2015

DIstribusi Binomial Negative

Contoh.

Sebuah obat diketahui efektif 60% dari total pemakaiannya. Kita ingin tahu probabilitasnya bahwa pasien ke lima yg sembuh memakai obat ini adalah pasien ke tujuh yg diberikan obat ini. Jadi jika S:sembuh dan G:gagal, salah satu kemungkinan urutan peristiwanya adalah SGSSSG**S.** Untuk urutan spt ini akan muncul dengan probabilitas : (0.6)(0.4) (0.6)(0.6) (0.6)(0.4) (0.6) = $(0.6)^5(0.4)^2$. Kita bisa mencari seluruh permutasi urutan S dan G, dengan kendala bahwa elemen ketujuh (terakhir) harus S yg kelima. Jadi banyaknya konfigurasi adalah sama dengan banyak cara mempartisi 6 elemen, menjadi 2 grup, yg terdiri dari 4S dan 2G, yang tak lain adalah kombinasi 6 elemen diambil 4! Jadi jika X adalah variabel random yg menyatakan bahwa no urutan outcome sukses kelima terjadi, maka

$$P(X = 7) = \binom{6}{4}(0.6)^5(0.4)^2$$

Ini adalah contoh distribusi probabilitas binomial negatif!

DIstribusi Binomial Negative

Keterangan:

Kita bisa memandang bahwa dalam kasus ini kita punya 6 elemen:

S_1 S_2 S_3 S_4 G_1 G_2. Pertama kita pikirkan banyaknya seluruh permutasi yg mungkin dari 6 elemen ini adalah 6! Tetapi karena sebenarnya $S_1=S_2=S_3=S_4=S$, maka banyak konfigurasi yg berbeda harus dibagi 4!

Misal : $S_1G_1S_2S_3G_2S_4 = S_2G_1S_1S_3G_2S_4 = S_2G_1S_2S_4G_2S_3 = \ldots.$ yaitu seluruh permutasi yg mungkin dari label 1,2,3,4 pada S

Demikian juga untuk G_1 dan $G_{2,}$, sehingga total konfigurasinya harus dibagi 2!,

Jadi total konfigurasi yg berbeda yg melibatkan 4S dan 2G adalah:

$$\frac{6!}{4!2!}$$ yg tak lain adalah kombinasi $$\binom{6}{4}$$

DIstribusi Binomial Negative

Jadi kita melakukan percobaan sebanyak x kali. Sebanyak k kali adalah “sukses” berarti (x-k) gagal. Misal probabilitas terjadinya “sukses” =p berarti probabilitas “gagal” q=1-p

Probabilitas untuk mendapatkan “sukses” pada percobaan ke-x, yg didahului oleh (k-1) sukses – berarti urutan ke-x tsb adalah “sukses” ke k, dan (X-k) “gagal”, dengan **urutan** “sukses” dan “gagal” tertentu adalah:

$$p^{k-1} q^{x-k} p = p^k q^{x-k}$$

Banyak kombinasi yg berbeda dari (x-1) elemen yg terdiri dari (k-1) “sukses” dan (x-k) “gagal” adalah kombinasi (x-1) diambil (k-1) elemen: $\binom{x-1}{k-1}$

Sehingga probabilitas bahwa dari x percobaan diakhiri dengan “sukses” yg ke-k kali adalah :

$$b^*(x;k,p) = \binom{x-1}{k-1} p^k q^{x-k}$$ (Distribusi Binomial Negatif)

DIstribusi Geometrik

Jika probabilitas sebuah “sukses” = p dan probabilitas “gagal” q=1-p, dan X adalah variabel random yg menyatakan **jumlah percobaan** yg diperlukan agar didapatkan “sukses” yg pertama kali, maka probabilitas g(x,p) = pq^{x-1}

INi disebut distribusi geometrik, yg tak lain adalah kasus khusus dari distribusi binomial negatif, dimana banyak sukses k=1 dan terjadi di akhir percobaan yg sebanyak X :

$$b^*(x;k=1,p)=\binom{x-1}{k-1}p^k q^{x-k}=\binom{x-1}{0}p\ q^{x-1}=\frac{(x-1)!}{(x-1)!0!}pq^{x-1}=pq^{x-1}$$

DIstribusi Geometrik

Contoh.

Dalam sebuah proses produksi diketahui, secara rata-rata 1 dari 100 hasil produksinya cacat. Berapakah probabilitasnya jikalau pada pengambilan kelima dari hasil produksinya dijumpai hasil produksi yg cacat pertama kali (jadi 4 yg pertama bagus)?

Jawab:

Ini adalah contoh distribusi geometrik, dengan jumlah percobaan $x=5$, probabilitas “sukses” yaitu produk cacat $p=0.01$, berarti probabilitas “gagal” $q=1-p = 1-0.01 =0.99$.

Jadi probabilitas mendapatkan hasil produk kelima adalah cacat yg pertama adalah :

$$g(x=5;p=0.01)= (0.01)(0.99)^{5-1} = 0.0096 = 0.96\%$$

DIstribusi Geometrik

Soal.

Pada jam-jam sibuk, bisa sulit sekali untuk sebuah panggilan telepon dilayani oleh jaringan telepon. Misalkan diketahui bahwa probabilitas untuk berhasil mendapatkan koneksi pada jam sibuk adalah 5%, berapakah probabilitasnya bahwa pada jam sibuk diperlukan 5 kali percobaan panggilan baru bisa mendapatkan koneksi?

DIstribusi Geometrik : Mean dan Variansi

Jika μ dan σ^2 menyatakan mean (rata-rata) dan variansi dari variabel random X yg memiliki distribusi geometrik, maka:

$$\mu = \frac{1}{p} \qquad \sigma^2 = \frac{1-p}{p^2}$$

DIstribusi Poisson – Variabel Random Poisson

Percobaan yg menghasilkan variabel random X yg menyatakan banyaknya outcome selama interval waktu tertentu atau dalam “area” atau “luas” tertentu dinamakan percobaan Poisson.

Contoh:

X : banyak panggilan telepon per jam

X : banyak hari-hari sekolah tutup karena bencana alam dalam setahun

X : banyaknya penundaan pertandingan bola karena hujan dalam semusim pertandingan

X : banyak tikus per hektare

X : banyaknya kesalahan ketik per halaman

Sifat Proses Poisson

1. Tidak punya memori atau ingatan, yaitu banyaknya outcome dalam satu interval waktu (atau daerah) tidak bergantung pada banyaknya outcome pada waktu atau daerah yg lain.
2. Probabilitas terjadinya 1 outcome dalam interval waktu (atau daerah) yg sangat pendek (kecil) sebanding dengan lama waktu interval waktu tsb (atau luas daerahnya). Dan tidak bergantung pada kejadian atau outcome di luar interval ini.
3. Probabilitas terjadinya lebih dari 1 outcome dalam interval waktu yg sangat pendek di (2) tsb sangat kecil atau bisa diabaikan.

Distribusi Poisson

X : variabel random Poission yg menyatakan banyaknya outcome selama percobaan.

μ : rata-rata banyak outcome = λt dimana t adalah lama intervalnya dan λ adalah laju terjadinya outcome.

Distribusi probabilitas dari variabel random Poisson X yg menyatakan banyaknya outcome dalam interval waktu tertentu t (atau daerah tertentu) dengan λ menyatakan laju terjadinya outcome persatuan waktu atau per satuan daerah diberikan oleh (tidak diturunkan!):

$$p(x;\lambda t) = \frac{e^{-\lambda t}(\lambda t)^x}{x!} \qquad \Rightarrow p(x;\mu) = \frac{e^{-\mu}(\mu)^x}{x!}$$

Selanjutnya ditabelkan distribusi kumulatif Poisson:

$$P(r;\mu) = \sum_{x=0}^{r} p(x;\mu)$$

Tabel Distribusi Poisson Kumulatif

μ = λt = mean out come

Table A.2 Poisson Probability Sums $\sum_{x=0}^{r} p(x;\mu)$

	μ					
r	0.1	0.2	0.3	0.4	0.5	0.6
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488
1	0.9953	0.9825	0.9631	0.9384	0.9098	0.8781
2	0.9998	0.9989	0.9964	[illegible]	0.9856	0.9769
3	1.0000	0.9999	0.9997	[illegible]	0.9982	0.9966
4		1.0000	1.0000	0.9999	0.9998	0.9996
5				1.0000	1.0000	1.0000
6						

$$P(r=3;\mu=0.4)=\sum_{x=0}^{3} p(x=3;\mu=0.4)=$$

Contoh

Dalam percobaan radioaktif, rata-rata jumlah cacahan radioaktif yg terekam di counter adalah 4 cacahan per mili detik. Berapakah probabilitasnya dalam 1 milidetik tertentu tercacah sebanyak 6 cacahan?

Jawab:

Rata-rata jumlah outcome per milidetik : $\mu = \lambda t = 4$

Probabilitas tercacah X=6 dalam 1 milidetik:

$$p(x=6;\mu=4)=\frac{e^{-\mu}(\mu)^{x}}{x!}=\frac{e^{-4}(4)^{6}}{6!}=0.1042$$

Atau dengan tabel Poisson:

$p(x=6;\mu=4)=P(r=6;\mu=4) - P(r=5;\mu=4) = 0.8893-0.7851=0.1042$

Soal

Rata-rata jumlah kedatangan kapal di suatu pelabuhan adalah 10 per hari. Fasilitas di pelabuhan tersebut di desain paling banyak mampu menangani 15 kapal per hari. Berapakah probabilitasnya bahwa pada hari tertentu kapal terpaksa tak bisa masuk pelabuhan karena telah penuh?

Mean dan Variansi

Distribusi Poisson $p(x;\lambda t)$ memiliki mean dan variansi = $\mu=\lambda t$

Contoh.

Dalam percobaan radioaktif, rata-rata jumlah cacahan radioaktif yg terekam di counter adalah 4 cacahan per mili detik. Berarti $\mu=4$ dan variansi $\sigma^2=4$ atau STD, $\sigma = 2$. Berarti menurut teorema Chebysev terdapat probabilitas 75% bahwa banyak cacahan yg datang dalam 1 milidetik akan terletak antara $\mu - 2\sigma$ dan $\mu + 2\sigma$ atau $4 – 2(2)$ dan $4+2(2)$ yaitu antara 0 s/d 8.

Hubungan Distribusi Poisson dan Binomial

Jika X adalah variabel random yg memiliki distribusi binomial b(x;n,p), maka jika jumlah percobaannya besar sekali $n \rightarrow \infty$ serta probabilitas untuk "sukses" p kecil sekali $p \rightarrow 0$, serta rata-ratanya yaitu $\mu = np$ maka dalam hal ini distribusi Binomial bisa diaproksimasi dengan distribusi Poisson.

$$b(x;n,p) \rightarrow p(x;\mu=np) \text{ untuk n besar, p kecil}$$

Contoh.

Probabilitas terjadinya kecelakaan di suatu hari di sebuah pabrik adalah 0.005.

(a) Berapakah probabilitasnya selama 400 hari tidak terjadi kecelakaan sama sekali?

(b) Berapakah probabilitasnya paling banyak 3 hari dengan dengan kecelakaan selama 400 hari tsb?

Hubungan Distribusi Poisson dan Binomial

Jawab:

X adalah variabel random binomial yg menyatakan banyak hari dengan kecelakaan (“sukses”) dengan probabilitas terjadinya kecelakaan dalam satu hari p=0.005.

a) Dalam 400 hari hanya 1 kecelakaan, berarti banyak percobaan n=400. Jadi ingin dihitung b(x=1;n=400,p=0.005). Karena n besar dan p kecil maka dapat dipergunakan aproksimasi distribusi Poisson dengan : rata-rata (mean) μ = np = 400*0.005= 2 kecelakaan dalam 400 hari, dan ingin diketahui probabilitasnya terjadi 1 kecelakaan dalam 400 hari tsb (X=1):

$$b(x=1;n=400,p=0.005) \rightarrow p(x=1;\ \mu=2)$$

$$p(x=1;\mu=2)=\frac{e^{-\mu}(\mu)^{x}}{x!}=\frac{e^{-2}(2)^{1}}{1!}=0.271$$

Hubungan Distribusi Poisson dan Binomial

Jawab:

b) Karena paling banyak jumlah hari dengan kecelakaan adalah 3 hari. Berarti ingin dihitung $P(x \leq 3)$, yaitu (dengan Tabel D. Poisson)

$$P\ (r=3;\ \mu=2) = 0.857$$

Jadi terdapat probabilitas 86% dalam 400 hari, akan terjadi kecelakaan 0, atau 1, 2 atau 3 paling banyak.

Soal

Dalam produksi peralatan dari gelas secara rata-rata dijumpai 1 gelas catat dari tiap 1000 produksi. Berapakah probabilitasnya dari 8000 gelas yg diproduksi jumlah gelas yg cacat kurang dari 7 buah?