Verjetnostni račun in statistika

Dodatno pojasnilo

8. januar 2014

Na predavanjih nismo do konca izpeljali zgleda, v katerem smo želeli pokazati, da lahko vsak primer z normalno porazdeljeno zvezno slučajno spremenljivko "prevedemo" na standardno normalno porazdelitev. Zato podajam natančno razlago.

Denimo, da je $X \sim N(\mu, \sigma^2)$ in naj bo $Y = \frac{X-\mu}{\sigma}$. Za $a = 1/\sigma$ in $b = -\mu/\sigma$ je torej $Y = aX + b$. Po trditvi 4.5 torej za gostoto slučajne spremenljivke Y velja $f_Y(x) = \frac{1}{a} \cdot f_X(\frac{x-b}{a})$. Ker je $\frac{x-b}{a} = \frac{x+\frac{\mu}{\sigma}}{\frac{1}{\sigma}} = \sigma x + \mu$, tako dobimo

$$f_Y(x) = \sigma \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{\sigma x+\mu-\mu}{\sigma}\right)^2} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}.$$

To pa pomeni, da je $Y \sim N(0, 1)$. Ta premislek torej pokaže, da v primeru, ko je $X \sim N(\mu, \sigma^2)$, velja $\frac{X-\mu}{\sigma} \sim N(0, 1)$ in tako lahko res vsako normalno porazdeljeno slučajno spremenljivko prevedemo na standardno normalno porazdelitev.

Oglejmo si še konkreten zgled uporabe tega dejstva. Denimo, da študentje pišejo kolokvij in naj bo X slučajna spremenljivka, ki meri število doseženih točk na kolokviju. Predpostavimo, da je X (približno) normalno porazdeljena in denimo, da velja $X \sim N(75, 15^2)$, to je, pričakovani povprečni rezultat je 75 točk, standardni odklon pa je 15. Izračunajmo tedaj verjetnost, da naključno izbrani študent doseže vsaj 90 točk.

Zanima nas torej verjetnost $P(X \geq 90) = 1 - P(X < 90)$. Dovolj je torej poznati vrednost $P(X < 90) = P(X \leq 90) = F_X(90)$. Po zgornjem za $Y = \frac{X-75}{15}$ velja $Y \sim N(0, 1)$. Ker je $X = 15Y + 75$, je po trditvi 4.5 $F_X(x) = F_Y(\frac{x-75}{15})$, torej je $F_X(90) = F_Y(1)$. Vrednosti porazdelitvene funkcije za standardno normalno porazdelitev so tabelirane. Iz takšne tabele potem odčitamo, da je $F_Y(1) \approx 0.8413$ in tako je $P(X \geq 90) \approx 0.1587$. Verjetnost, da bo naključno izbrani študent pisal vsaj 90 točk, je potemtakem slabih 16 procentov. Seveda lahko do tega rezultata pridemo tudi direktno, brez uporabe trditve 4.5. Tedaj je potrebno pač izračunati

$$P(X \leq 90) = P(X - 75 \leq 15) = P(\frac{X-75}{15} \leq 1) = P(Y \leq 1) = F_Y(1).$$

Izračunajmo še verjetnost, da študent ne doseže niti 50 točk. Tokrat nas seveda zanima verjetnost $P(X < 50) = P(X \leq 50) = F_X(50)$. Kot prej dobimo $F_X(50) = F_Y(\frac{-25}{15}) = F_Y(\frac{-5}{3})$. Te vrednosti v običajnih tabelah ne bomo našli. A ker je gostota f_Y soda funkcija (simetrična na y-os), je dovolj poznati vrednost $F_Y(\frac{5}{3})$, saj je $F_Y(\frac{-5}{3}) = 1 - F_Y(\frac{5}{3})$. Ker je $F_Y(\frac{5}{3}) \approx 0.9525$, je tako $F_X(50) \approx 0.0475$, torej je verjetnost, da študent ni zbral niti 50 točk, slabih 5 procentov.