Egenskaper hos normal-, χ^2-, t- och F-fördelningarna

Martin Singull

21 oktober 2011

Definition 1. *En sv. X kallas normalfördelad om den har täthetsfunktion*

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

Sats 1. *Om $X \sim N(\mu, \sigma)$ så gäller att* $Z = \dfrac{X-\mu}{\sigma} \sim N(0,1)$.

Sats 2. *Om de stokastiska variablerna $X_1, \ldots, X_n$ är oberoende, om $X_i \sim N(\mu_i, \sigma_i)$ och om $a_1, \ldots, a_n$ och b är konstanter, så gäller att den sv.*

$$Y = \sum_1^n a_i X_i + b \sim N\left(\sum_1^n a_i\mu_i + b, \sqrt{\sum_1^n a_i^2\sigma_i^2}\right).$$

Notera att parametrarna är väntevärdet och standardavvikelsen.

Definition 2. *En sv. X med täthetsfunktion av typen*

$$f(x) = k \cdot x^{f/2-1} \cdot e^{-x/2} \quad \textit{för} \quad x > 0$$

kallas χ^2-fördelad med f frihetsgrader. Man skriver $X \sim \chi^2(f)$.

Man får χ^2-fördelade stokastiska variabler genom att summera kvadrater på oberoende $N(0,1)$-variabler. Man kan nämligen visa

Sats 3. *Om $Z_1, \ldots, Z_n$ är oberoende och $N(0,1)$, så gäller att*

$$Z_1^2 \sim \chi^2(1) \quad \textit{och} \quad \sum_1^n Z_i^2 \sim \chi^2(n)\ .$$

Följdsats 1. *Om $X \sim \chi^2(f)$ så gäller att $E(X) = f$.*

Följdsats 2. *Om X och Y är oberoende, $X \sim \chi^2(f_1)$ och $Y \sim \chi^2(f_2)$, så gäller att $X + Y \sim \chi^2(f_1 + f_2)$.*

Vi har också följande sats, som vi kommer att utnyttja mycket.

Sats 4. *Om $X_1, \ldots, X_n$ är oberoende och $X_i \sim N(\mu, \sigma)$, så gäller att*

(a) $\sum_1^n \left(\dfrac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(n)$

(b) $\dfrac{(n-1)S^2}{\sigma^2} = \dfrac{1}{\sigma^2}\sum_1^n \left(X_i - \bar{X}\right)^2 \sim \chi^2(n-1)$

(c) $\bar{X} \sim N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right)$

(d) $\bar{X}$ och S^2 är oberoende stokastiska variabler.

I fall (a) ser man ganska lätt att det är en summa av kvadrater på oberoende $N(0,1)$-variabler; (c) har vi redan visat. Genom att uynyttja teorin för flerdimensionell normalfördelning kan man visa (b) och (d).

Definition 3. *En sv. X kallas t-fördelad med f frihetsgrader om den har en täthetsfunktion av typen*

$$f(x) = c \cdot \frac{1}{(1 + x^2/f)^{(f+1)/2}}\ .$$

Notera att täthetsfunktionen för en t-fördelad stokastisk variabel har den trevliga symmetriegenskapen $f(-x) = f(x)$.

Man kan också visa att $t(f)$-fördelning konvergerar mot $N(0,1)$-fördelning då $f \to \infty$.

Hur t-fördelningen uppstår framgår av följande sats. Lägg märke till att frihetsgraden f kommer från χ^2-variabeln i nämnaren.

Sats 5. *(Gosset/Student) Om X och Y är oberoende, $X \sim N(0,1)$ och $Y \sim \chi^2(f)$, så gäller att*

$$\frac{X}{\sqrt{Y/f}} \sim t(f)\ .$$

Gossets sats kommer vi att använda ofta då vi tar fram hjälpvariabler för att konstruera konfidensintervall för väntevärdesparametrar i normalfördelningen.

Låt t.ex. $X_1, \ldots, X_n$ vara oberoende och $N(\mu, \sigma)$. Genom att kombinera resultaten i de två senaste satserna får vi

$$\frac{(\bar{X}-\mu)/(\sigma/\sqrt{n})}{\sqrt{\frac{1}{\sigma^2}(n-1)S^2/(n-1)}} = \frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)\ .$$

Slutligen ska vi se hur F-fördelningen uppstår.

Sats 6. *Om Y_1 och Y_2 är oberoende, $Y_1 \sim \chi^2(r_1)$ och $Y_2 \sim \chi^2(r_2)$, så gäller att*

$$V = \frac{Y_1/r_1}{Y_2/r_2} \sim F(r_1, r_2)$$

dvs. V är F-fördelad med r_1 och r_2 frihetsgrader.

Vi ser att $\dfrac{1}{V} \sim F(r_2, r_1)$.

F-fördelningen använder vi i den här kursen vid jämförelse av varianser och i samband med vissa test i variansanalys.

Anm. Om man tar den stokastiska variabeln i Gossets sats i kvadrat så får man

$$\frac{X^2}{Y/f} \sim F(1, f)\ .$$

Kvadraten på en $t(f)$-variabel är alltså en stokastisk variabel som är $F(1, f)$.

OBS! Montgomery använder beteckningen $N(\mu, \sigma^2)$, där σ^2 = variansen, men vi fortsätter att skriva $N(\mu, \sigma)$, där σ = standardavvikelsen, precis som i grundkursen.