Review of Probability

Econometrics

Péter Elek

Review: Concept of Probability

- Events
- Rules of probability (axioms)
- Conditional probability
- Independence of events
- Bayes's rule

- See e.g., Cunningham 2.1-2.6.

Random Variables (r.v.)

- Informal definition: a variable whose value is determined by the outcome of a random experiment
- Notation: X = r.v., x = realized value
- Example: flipping a coin 10 times and counting the number of heads
- Types: discrete and continuous
  - Discrete: e.g., coin toss
  - Continuous: e.g., uniform distribution on [0,1]

Probability Distribution of Discrete Random Variables

- Probability distribution: $\Pr(X = x_j) = p_j$, with $\sum p_j = 1$
  - Die: $\Pr(X = i) = 1/6$ $(i = 1, \ldots, 6)$
  - Coin: $\Pr(X = \text{heads}) = \Pr(X = \text{tails}) = 1/2$

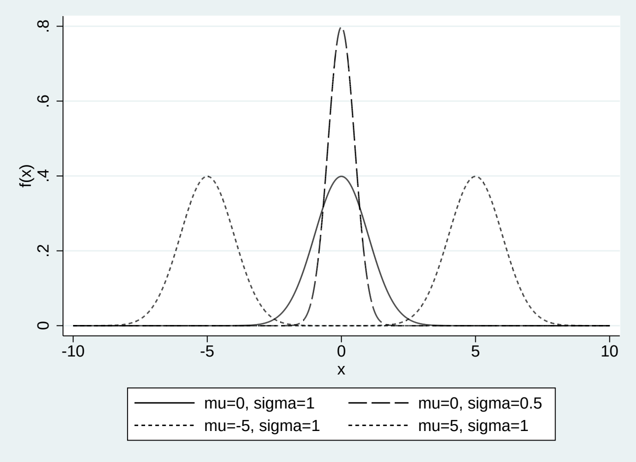
Cumulative Distribution Function (CDF)

- $F(x) = \Pr(X \leq x)$
- Properties:
  - $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$
  - Non-decreasing
  - $\Pr(X > a) = 1 - F(a)$
  - $\Pr(a < X \leq b) = F(b) - F(a)$

Probability Density Function (PDF) of Continuous Random Variables

- PDF: $f(x) = F'(x)$
- $\Pr(a < X \leq b) = \int_a^b f(x)dx$
- Note: $\Pr(X = x) = 0$, but $f(x)$ indicates "typical" values: $\Pr(x < X \leq x + dx) \approx f(x)dx$
- Properties:
  - $f(x) \geq 0$
  - $\int_{-\infty}^{\infty} f(x)dx = 1$

$N(\mu, \sigma^2)$ Normal PDF: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}$



Conditional Distribution and Independence

- Conditional distribution of Y given $X = x$: $f_{Y|X}(y|x)$ (continuous) or $\Pr(Y = y_j|X = x_i)$ (discrete)
- X and Y are independent if the conditional distribution does not depend on the other variable

Expected Value

- Discrete: $E(X) = \sum p_i x_i$
- Continuous: $E(X) = \int x f(x) dx$
- Properties:
  - $E(c) = c$
  - $E(aX + b) = aE(X) + b$
  - $E(a_1 X_1 + \cdots + a_k X_k) = a_1 E(X_1) + \cdots + a_k E(X_k)$
- Minimizes mean squared deviation: $b = E(X)$ minimizes $E[(X - b)^2]$

Variance and Standard Deviation

- $Var(X) = E[(X - E(X))^2] = E(X^2) - (E(X))^2$
- Standard deviation: $\sigma_X = \sqrt{Var(X)}$
- Properties:
  - $Var(aX + b) = a^2 Var(X)$
  - $\sigma_{aX+b} = |a|\sigma_X$
  - $Var(c) = 0$
- Let $E(X) = \mu$ and $Var(X) = \sigma^2$. Then $Z = \frac{X-\mu}{\sigma}$ is the standardized r.v., for which $E(Z) = 0$ and $\sigma(Z) = 1$.

Covariance and Correlation

- Covariance: $Cov(X, Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y)$
- Correlation: $Corr(X, Y) = Cov(X, Y)/(\sigma_X \sigma_Y)$
- Properties:
  - $-1 \leq Corr(X, Y) \leq 1$
  - independence $\implies$ $Cov(X, Y) = 0$ (but the converse is not necessarily true)
  - $Cov(aX + b, cY + d) = ac \cdot Cov(X, Y)$
  - $Corr(aX + b, cY + d) = sign(ac) \cdot Corr(X, Y)$ (invariant to scaling)

Variance of Linear Combinations

- $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2abCov(X, Y)$
- If uncorrelated: $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y)$

Conditional Expectation

- $E(Y|X = x)$: expected value of Y given $X = x$
- Discrete: $E(Y|X = x) = \sum y_j \Pr(Y = y_j|X = x)$
- Continuous: $E(Y|X = x) = \int y f_{Y|X}(y|x) dy$
- Example: X = education, Y = monthly wage,
  $E(Y|X = \text{university}) > E(Y|X = \text{8th grade})$
- Notation: $E(Y|X)$ is the r.v. that takes the value $E(Y|X = x)$ if $X = x$.

Importance of Normal Distribution

- Many variables in real life (e.g., male or female height) are approximately normally distributed
- Many estimators are approximately normal in large samples
- Derived distributions: lognormal, χ^2, t, F distributions

Properties of Normal Distribution

- $X \sim N(\mu, \sigma^2) \implies aX + b \sim N(a\mu + b, (|a|\sigma)^2)$
- Standardization: $(X - \mu)/\sigma \sim N(0, 1)$ (standard normal distribution)
- Sum of independent normals: $X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$

Calculating Normal Probabilities

- $\Phi(z)$: CDF of $N(0,1)$
- $\Pr(a < X \leq b) = \Phi((b-\mu)/\sigma) - \Phi((a-\mu)/\sigma)$
- Example: $\Pr(|(X-\mu)/\sigma| \leq 1.96) = 0.95$

Exercise

Suppose that the heights of men (m) and women (w) are normally distributed with $\mu_m = 176$, $\mu_w = 164$, $\sigma = 7$.

a. Find the probability that a randomly selected woman is taller than 176 cm.
b. Find the probability that a randomly selected woman is taller than a randomly selected man.
c. If spouses' heights have a correlation of 0.3, then find the probability that the wife is taller than the husband.

Material

- Cunningham 2.1-2.10
- Wooldridge Appendix B (excl. χ^2, t, F distribution)