

Lärmål Sannolikhet, statistik och risk 2015

Johan Jonasson ^{*†‡}

December 2015

Följande begrepp och metoder ska behärskas väl, kunna förklaras och tillämpas. Direkta bevis av satser från kursen kommer inte på tentan, men bevisen ska ändå behärskas, eftersom delar av dem mycket väl kan komma att ingå som del av lösningar till tentamensuppgifter.

Övningsuppgifterna nedan är rekommenderade övningar för respektive moment. Naturligtvis kan många övningsuppgifter behandla flera olika moment.

- Behärska och kunna förklara begreppen utfallsrum, utfall och händelse. Att kunna tolka händelser, både formellt som delmängder till utfallsrummet, dels informellt ”i text” och kunna fritt röra sig mellan de båda. Övningar i kap 1: 1, 3, 5.
- Definitionen av sannolikhetsmått och dess omedelbara konsekvenser. Övningar i kap 1: 7, 10, 11, 12, 13.
- Kombinatorik och det klassiska sannolikhetsmättet. Här ingår multiplikationsprincipen och att kunna beräkna på hur många sätt man kan välja k element ur en mängd med n element, både med och utan återläggning och med och utan hänsyn till ordning. Binomialkoefficienter och deras kombinatoriska betydelse. Övningar i kap 1: 19, 22, 24, 27, 30, 34, 35.
- Definitionen och den intuitiva betydelsen av betingad sannolikhet, man stryker en del av utfallsrummet. Övningar i kap 1: 37, 39, 42, 49, 50.
- Oberoende händelser, även för fler än två händelser. Förstå att oberoende ofta kommer in som en del av en modell. Övningar i kap 1: 40, 44, 45, 52, 53, 54, 58, 62, 63, 67, 70.

^{*}Chalmers University of Technology

[†]Göteborg University

[‡]jonasson@chalmers.se

- Totala sannolikhetslagen och Bayes formel (att vända på betingning). Övningar i kap 1: 72, 76, 78, 80, 87, 88, 89, 90, 92, 95, 98, 100, 103.
- Rekursivt tänkande med hjälp av totala sannolikhetslagen. Övningar i kap 1: 105, 108.
- Definitionen av en stokastisk variabel. Fördelningen av en stokastisk variabel och hur den bestäms av fördelningsfunktionen $F(x) = \mathbb{P}(X \leq x)$. Frekvensfunktionen (pmf), $p(x_k) = \mathbb{P}(X = x_k)$ av en diskret stokastisk variabel och täthetsfunktionen (pdf) $f(x) = F'(x)$ av en kontinuerlig stokastisk variabel och alla samband mellan dessa och fördelningsfunktionen. Övningar i kap 2: 1, 2, 3, 6, 9, 10, 13, 14, 16, 17, 18, 22.
- Väntevärde av en stokastisk variabel, inklusive dess tolkning som medelvärde i det långa loppet och som tyngdpunkt hos tätheten/frekvensfunktionen. Formeln $\mathbb{E}[X] = \int_0^\infty \mathbb{P}(X > x)dx$ för en positiv kontinuerlig stokastisk variabel och $\mathbb{E}[X] = \sum_{k=1}^\infty \mathbb{P}(X \geq k)$ för en positiv diskret sv. Variansen av en stokastisk variabel och Steiners formel $\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$. Övningar i kap 2: 24, 26, 27, 28, 31, 33, 34, 38, 39, 42, 45, 47.
- Funktioner av stokastiska variabler: att via fördelningsfunktion kunna beräkna fördelningen av en funktion av en stokastisk variabel. Den omedvetne statistikerns lag (Proposition 2.12) för beräkning av väntevärde av en funktion av en stokastisk variabel. Övningar i kap 2: 40, 41, 43.
- Likformig fördelning: täthet, väntevärde, varians. Övningar i kap 2: 19, 20.
- Binomialfördelning: frekvensfunktion, väntevärde, varians, att skriva som summa av indikatorer. Övningar i kap 2: 49, 50, 52, 54, 56, 57, 58.
- Geometrisk fördelning: frekvensfunktion, väntevärde, varians, samband med binomialfördelning (om X är geometrisk(p) och Y är $\text{Bin}(n, p)$, så är $\mathbb{P}(X > n) = \mathbb{P}(Y = 0)$), samband med exponentialfördelning (om X är exponentialfördelad är X avrundad till närmast större heltal geometriskt fördelad). Övningar i kap 2: 59, 61, 62.
- Poissonfördelning: frekvensfunktion, väntevärde, varians, samband med exponentialfördelning (se också Poissonprocessen). Övningar i kap 2: 64, 65.
- Exponentialfördelning: täthet, väntevärde, varians, glömskeegenskapen (se också Poissonprocessen). Övningar i kap 2: 70, 71, 72, 74, 76.

- Normalfördelning: täthet, väntevärde, varians, standardisering till $N(0, 1)$ och användning av tabeller för Φ . Percentilbegreppet. Övningar i kap 2: 77, 78, 79, 80, 82, 83, 85, 86, 88.
- (Endast TM) Gamma-, χ^2 -, Weibull-, Gumbel- och lognormalfördelningarna. Övningar i kap 2: 92, 93, 94, 96, 97, 98 och i Matlabhäftet.
- Felintensitetsfunktionen av en kontinuerlig stokastisk variabel och hur man beräknar tätheten utifrån den. Exponentialfördelningens speciella roll som den fördelning som har konstant felintensitet. Övningar i kap 2: 108, 110, 112, 114 och i Matlabhäftet.
- Flerdimensionella fördelningar. Definition av flerdimensionell fördelningsfunktion, frekvensfunktion och täthet och sambanden mellan dessa. Marginalfördelningar för de enskilda ingående stokastiska variablerna. Övningar i kap 3: 1, 4, 7, 9, 11, 13, 15, 16, 18, 19, 20.
- Betingade fördelningar (betingad fördelningsfunktion och täthet/frekvensfunktion). Varför ser den betingade tätheten ut som den gör? Oberoende stokastiska variabler: definition och hur oberoende kan karaktäriseras av faktorisering av fördelningsfunktion och faktorisering av täthet/frekvensfunktion. Övningar i kap 3: 22, 23, 24, 25, 27, 30, 32, 33, 35, 36, 37, 39, 41.
- Funktioner av flerdimensionella stokastiska variabler. Den flerdimensionella varianten av den omedvetne statistikerns lag (Proposition 3.11). Speciellt viktiga resultat av denna är (1) väntevärdet av en summa är alltid summan av väntevärdena, (2) för oberoende variabler är väntevärdet av en produkt lika med produkten av väntevärdena och variansen av en summa är summan av varianserna. Övningar i kap 3: 43, 46, 47, 48, 50, 51, 52, 54, 55, 58, 61, 63, 65.
- Betingade väntevärden och den totala sannolikhetslagen för väntevärden (Proposition 3.17 & Corollary 3.5). Övningar i kap 3: 68, 72, 73, 74, 76, 77, 81, 82, 83.
- Kovarians, korrelation och korrelationskoefficient. Varians är ett specialfall kovarians. X och Y oberoende medför att $\text{Cov}[X, Y] = 0$, men omvändningen gäller inte allmänt. Schwarz olikhet. Bästa linjära prediktor. Övningar i kap 3: 85, 87, 89, 92, 93, 94, 96, 98, 99.
- Faltning (convolution) av oberoende stokastiska variabler. Speciellt viktigt är att summor av oberoende Poissonfördelade sv är Poisson och att summor

av oberoende normalfördelade sv är normalfördelade. Övningar i kap 3: 125, 126, 128, 129, 130, 131.

- Poissonprocessen: definition och uttunnings- och sammanvägningsegenskaperna och sambandet med exponentialfördelningen. Övningar i kap 3: 150, 152, 154, 157, 158 och i Matlabhäftet.
- Momentgenererande funktion (mgf). Definitionen och hur momenten beräknas utifrån mgf. Övningar i kap 3: 144, 145, 148.
- (Endast TM) Simulering och transformering mellan olika fördelningar. Övningar i kap 5: 1, 2, 3, 7, 11, 14, 16 och i Matlabhäftet.
- Stora talens lag och dess bevis via Chebyshevs olikhet. Övningar i kap 4: 1, 3, 5.
- Centrala gränsvärdessatsen och bevisskiss med momentgenererande funktion. Specialfallen med tillämpning på binomialfördelning och Poissonfördelning. Tillämpningar och en medvetenhet om dess begränsningar: se upp med svansarna! Övningar i kap 4: 8, 9, 11, 12, 13, 15.
- Väntevärdesriktighet och konsistens hos punktskattningar. Medelvärdet $\bar{X}$ är en väntevärdesriktig skattning av väntevärdet och stickprovsvariansen s^2 för variansen. Övningar i kap 6: 1, 3, 5, 7, 10, 11, 14, 15.
- Konfidensintervall: allmän princip och de viktiga specialfallen med konfidensintervall för väntevärde i normalfördelningen, både med känd och okänd varians, konfidensintervall för p i binomialfördelningen och för skillnaden $p_1 - p_2$ för två oberoende binomialfördelningar. Övningar i kap 6: 17, 18, 19, 20, 21, 23, 27, 28, 30 och i kap 7: 9, 13.
- ML-principen (maximum likelihood principle) för punktskattningar. Vad vill principen säga? Övningar i kap 6: 34, 35, 37, 39.
- Statistiska tester: allmän princip och korrespondensen med konfidensintervall. Övningar i kap 6: 43, 44, 49.
- De olika normalfördelningstesterna, för ett och två stickprov, med känd eller okänd varians. Test av p i binomialfördelningen och av $p_1 - p_2$ för två oberoende binomialfördelningar. Övningar i kap 7: 15, 16, 18, 19, 20, 25, 26, 27, 29, 32 och i Matlabhäftet.
- P-värden och styrkor för ett test. Övningar i kap 6: 59, 60, 62 och i Matlabhäftet.

- Förstå problemen med multipeltestning och tjuvkikande på data. Övningar i kap 6: 53, 55.
- (Endast TM) Variansanalys och F -fördelningen. Övningar i kap 7: 34 och i Matlabhäftet.
- Linjär regression: modell, ML-skattningar, konfidensintervall och prediktion. Övningar i kap 7: 39, 41, 45, 46 och i Matlabhäftet.
- Bayesiansk statistik: vara bekant med tankesättet och förstå skillnaden gentemot den frekventistiska traditionen. Förstå fördelar och nackdelar. Övningar i kap 6: 76, 77, 78, 79, 80, 82, 83.
- (Endast TM) Ickeparametriska statistiska metoder: teckentestet och de två Wilcoxontesterna. Övningar i kap 6: 88, 89, 91, 97, 99, 101 och i Matlabhäftet.
- (Endast TM) Bootstrap-simulering: grundfilosofi och kännedom om möjliga fällor.