1 Descrizione sommaria degli argomenti svolti a lezione

Lezione 1 (25/2, 2 ore). Introduzione al corso (testo: S. Ross, materiale complementare in rete alle pagine di Gubinelli e Flandoli, primo compitino previsto per la settimana del 19 Aprile). Capitolo 3: oggetti del calcolo delle probabilità: esiti, spazio degli esiti, eventi ed operazioni su di essi, probabilità e sue regole. Probabilità condizionale, formula di fattorizzazione,

Lezione 2 (26/2, 1 ora). Formula di Bayes. Soluzione dell'es. 1 del compito d'esame del 4/6/02 (parti (i) e (ii)) (nota: prima è stata data una risoluzione intuitiva per via grafica, poi quella rigorosa con le probabilità condizionali e sue formule).

Lezione 3 (28/2, 2 ore) Spazi equiprobabili, principio di enumerazione. Esercizi n. 19 e 18 del Ross. Es. n. 29 su fattorizzazione e Bayes. Assegnati es. 31 (media difficoltà) e 27 (difficile) del Ross, per la lezione successiva.

Lezione 4 (3/3, ridotta ad un'ora per assemblea di Facoltà). Indipendenza, esempi elementari su sistemi in serie ed in parallelo. Esercizi assegnati: 6,7,8,9 (nota: la soluzione in rete del 9 è sbagliata: è $1-$ il valore in rete) della lista "Alcuni esercizi (parzialmente risolti)" della pag. web di Gubinelli, 37 e 41 del Ross.

Lezione 5 (4/3, 1 ora). Soluzione degli esercizi n. 27, 31, 41.

Lezione 6 (6/3, 2 ore). Da ora si inizia lo studio congiunto dei capitoli 4 e 5 sulle variabili aleatorie. Generalità ed esempi elementari. V.a. discrete, massa di probabilità (legge, distribuzione), v.a. di Bernoulli. Somma di n v.a. di Bernoulli di parametro p, indipendenti: calcolo della sua massa, definizione di v.a. binomiale. In questo calcolo è stato inserito il conteggio di tutti i modi in cui si possono porre k asterischi in n caselle, che produce il numero $\binom{n}{k}$ (par. 3.5.1).

Lezione 7 (10/3, 2 ore). V.a. discrete: grafico della massa di probabilità; funzione di distribuzione $F(x)$ e sua equivalenza (come contenuto di informazione) con la massa; es. di Bernoulli. Valor medio: definizione e motivazione (collegamento con la media aritmetica, tramite le frequenze relative, p. 112). Esempio della Bernoulli. Esercizi per casa: 1) calcolo di $F(x)$ per la v.a. X tale che $P(X = 2) = \frac{1}{2}$, $P(X = 0) = \frac{1}{4}$, $P(X = -2) = \frac{1}{4}$. 2) Trovare una v.a. X con $E[X] = -1$. 3) (difficile) Calcolare la media di una binomiale. 4) Completare l'esercizio 1 del compito del 4/6/02 e svolgere la parte (ii) del secondo esercizio (solo un calcolo della media).

Lezione 8 (11/3, 1 ora). Proprietà del valor medio: linearità e positività (senza dimostrazione); $E[XY] = E[X]E[Y]$ per X, Y indipendenti. Vettori aleatori (X, Y) e loro massa di probabilità congiunta $p(x, y)$. Calcolo delle masse marginali $p_X(x)$ e $p_Y(y)$ a partire dalla congiunta; calcolo della congiunta a partire dalle marginali nel caso di v.a. indipendenti. Dimostrazione di $E[XY] = E[X]E[Y]$ per v.a. discrete, usando la massa congiunta. Esercizio: calcolo della media di una binomiale $B(n, p)$.

Lezione 9 (17/3, 2 ore). Riassunto su v.a. discrete e loro valor medio. Valor medio di una trasformazione: $E[f(X)]$. Varianza $Var[X]$. Uguaglianza con $E[X^2] - E[X]^2$. Osservazione sulla diversità, in generale, tra $E[X^2]$ e $E[X]^2$. Due esercizi di calcolo della varianza (per una Bernoulli ed una simile). Teorema: $Var[X+Y] = Var[X] + Var[Y]$ per X, Y indipendenti, con dimostrazione. Esercizio: varianza di una binomiale $B(n, p)$. Interpretazione grafica della varianza; altro nome: scarto quadratico medio. Definizione di deviazione standard. Per casa: es. 1,2,3,7 del compito del 16/7/02.

Lezione 10 (18/3, 1 ora). Formula per $Var[X+Y]$ in generale e definizione di $Cov(X, Y)$. Uguaglianza con $E[XY] - \mu_X \mu_Y$. Definizione di coefficiente di correlazione $\rho(X, Y)$. Sua indipendenza dall'unità di misura: $\rho(\alpha X, \beta Y) = \rho(X, Y)$. Durante la dimostrazione: $Var[\alpha X] = \alpha^2 Var[X]$.

Lezione 11 (20/3, 2 ore). Riassunto di tutte le definizioni e proprietà relative ai valori medi. Ulteriori proprietà (tutte dimostrate): $Var[X + \alpha] = Var[X]$; $E\left[\frac{X-\mu}{\sigma}\right] = 0$; $Var\left[\frac{X-\mu}{\sigma}\right] = 1$.

Definizione di funzione generatrice dei momenti $\phi(t)$. Suo calcolo per la Bernoulli e binomiale e teorema sulla generatrice della somma di v.a. indipendenti (dimostrato). Legame con i momenti attraverso le derivate ed esemplificazione per la Bernoulli.

Esercizi per casa: 1) calcolare media e varianza di una binomiale usando la sua generatrice. 2) Calcolare media e varianza di $\frac{X_1+...+X_n}{n}$ quando $X_1, ..., X_n$ sono indipendenti, tutte con media μ e deviazione standard σ.

Lezione 12 (25/3, 2 ore). Soluzione degli esercizi 1,2,3 del compito del 16/7/02. Funzione generatrice della binomiale. Calcolo di media e varianza della binomiale usando la generatrice. Grafico della binomiale, simmetria per $p = \frac{1}{2}$ e riflessione dei grafici sotto la trasformazione $p \to 1 - p$. Concentrazione per n grande.

Lezione 13 (26/3, 1 ora). Teorema degli eventi rari (convergenza della binomiale alla Poisson), con dimostrazione. Esercizio: quante linee deve avere a disposizione una centrale telefonica con 1000 abbonati, ciascuno che si connette in un minuto tipo con probabilità $\frac{1}{50}$, per servire tutte le richieste di quel minuto tipo con probabilità 0.99?

Lezione 14 (27/3, 2 ore): sospesa per sciopero.

Lezione 15 (28/3, 2 ore). Variabili aleatorie continue: definizione, definizione di densità di probabilità, calcolo di $P(X \in A)$ per diversi insiemi A, prime osservazioni (es. $P(X = x) = 0$ per ogni x). Calcolo di $F(x)$. Esercizio: data $X = C$ per $x \in [0, 2]$ e zero altrimenti, calcolare C, calc. $P(X \in [-1, 1])$, calc. $F(x)$. Definizione di v.a. uniforme.

Definizione di valor medio, giustificazione intuitiva (usando $P(X \in [x, x + \varepsilon]) \sim \varepsilon f(x)$). Esempio: $E[X]$ per X uniforme su $[0, 1]$.

V.a. esponenziali di parametro λ: densità, suo studio di funzione (differenze al variare di λ), formula per $P(X \geq t)$. Teorema su $X = \min(X_1, ..., X_n)$ con X_i esponenziali indipendenti. Esempio di un servizio con n sportelli, calcolo di $P(X \geq t)$.

Lezione 16 (31/3, 2 ore). Soluzione dell'esercizio su media e varianza di $X = \frac{X_1 + ... + X_n}{n}$. $E[X] = \mu$, $Var[X] = \frac{\sigma 2}{n}$, interpretazione grafica (in vista della statistica).

V.a. esponenziali: $\int_a^b e^{-\lambda x} dx = \frac{e^{-\lambda a} - e^{-\lambda b}}{\lambda}$ e sue conseguenze. Dimostrazione del teorema sul minimo di v.a. esponenziali. Teorema sulla mancanza di memoria dell'esponenziale: enunciato, dimostrazione, interpretazione pratica. Discussione sul realismo dell'esponenziale nelle applicazioni.

V.a. di Poisson: definizione, richiamo sul legame con le binomiali e conseguente interpretazione come numero di eventi rari. Legame col numero di arrivi esponenziali.

Verifica della proprietà $\sum_k p_k = 1$ per binomiali e Poisson.

Lezione 17 (1/4, 1 ora). Esercizi 32, 39, 40, 24 della lista.

Lezione 18 (3/4, 2 ore). Definizione di gaussiana canonica (standard), suo grafico (tramite studio di funzione). Definizione di $\Phi(x)$ (funzione di distribuzione gaussiana canonica), suo calcolo tramite tabelle, $\Phi(-x) = 1 - \Phi(x)$. Definizione di quantile q_α (gaussiano canonico), suo calcolo tramite le tabelle. Definizione di z_α, relazione $z_\alpha = q_{1-\alpha}$.

Definizione di gaussiana generica $N(\mu, \sigma^2)$, suo grafico, ruolo grafico dei due parametri. Enunciato (senza dimostrazione) delle proprietà $\int_{-\infty}^{+\infty} f(x)dx = 1$, $E[X] = \mu$, $Var[X] = \sigma^2$, $\phi(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$, $F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$. Calcolo di $E[X]$ da $\phi(t)$.

Teorema su: $X \sim N(\mu, \sigma^2)$ implica $\alpha X + \beta \sim N(\alpha\mu + \beta, \alpha^2\sigma^2)$; dimostrazione con la generatrice ed anche dimostrazione con i valori medi del fatto che $\alpha\mu + \beta$ e $\alpha^2\sigma^2$ sono i parametri giusti.

Concetto di classe di v.a. autoriproducente (la somma di due indipendenti resta nella classe). Le Bernoulli non lo sono. Verifica in due modi che le $B(n,p)$ (n variabile) lo sono. Enunciato del fatto che lo sono anche le Poisson e le Gaussiane.

V.a. di Poisson: calcolo di $\phi(t)$, calcolo di $E[X]$ da $\phi(t)$, enunciato della formula $Var[X] = \lambda$.

Lezione 19 (7/4, 2 ore). Esercizi: Compito del 4/6/02, esercizio 2. Compito del 1/9/02, esercizio 1. Compito del 26/6/02, esercizio 2 con preliminare sulla funzione generatrice della distribuzione esponenziale. Compito del 22/7/03, esercizio 2.

Lezione 20 (17/4, 2 ore). Esercizi di preparazione al compitino.

Lezione 21 (21/4, 2 ore). Capitolo 6: considerazioni generali sulla statistica, sul senso del campionamento e sulla modellizzazione rigorosa tramite il concetto di campione di rango N di una distribuzione incognita.

Media campionaria o empirica, definizione, $E\left[\overline{X}\right] = \mu$, $Var\left[\overline{X}\right] = \frac{\sigma^2}{N}$. Accenno al fatto

che diminuendo la varianza ci si attende che la media empirica si concentri intorno a μ e quindi "stimi" μ.

Teorema del limite centrale, enunciato come convergenza della legge della somma (opportunamete riscalata e traslata) alla gaussiana.

Interpretazione come possibile tecnica di approssimazione di una somma di v.a. i.i.d.; Esercizio 6.3.1.

Approssimazione della Binomiale e confronto con il teorema degli eventi rari (diversi regimi di applicabilità).

Distribuzione approssimata della media campionaria usando il TLC.

Varianza campionaria, definizione e proprieta $E[S^2] = \sigma^2$, con dimostrazione. Cenno al fatto che è uno stimatore per σ^2.

Lezione 22 (24/4, 2 ore). Stima dei parametri (Cap. 7): i concetti di stima puntuale ed intervallare. Richiamo su alcune ragioni per cui si ritiene che $\overline{X}$ sia una buona stima puntuale di μ.

Richiamo sulle definizioni di q_β, z_α, il legame $z_\alpha = q_{1-\alpha}$, il significato geometrico di $z_{\frac{\alpha}{2}}$: $P\left(-z_{\frac{\alpha}{2}} \leq X \leq z_{\frac{\alpha}{2}}\right) = 1-\alpha$ se $X \sim N(0,1)$. Richiamo su $E\left[\overline{X}\right] = \mu$, $Var\left[\overline{X}\right] = \frac{\sigma^2}{N}$e gaussianità di $\overline{X}$ se si parte da gaussiane. Conseguenza: $P\left(-z_{\frac{\alpha}{2}} \leq \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \leq z_{\frac{\alpha}{2}}\right) = 1-\alpha$. Deduzione del teorema sull'intervallo di confidenza: $P\left(\overline{X} - \frac{\sigma z_{\frac{\alpha}{2}}}{\sqrt{n}} \leq \mu \leq \overline{X} + \frac{\sigma z_{\frac{\alpha}{2}}}{\sqrt{n}}\right) = 1-\alpha$. Sue riscritture, come $\mu = \overline{X} \pm \frac{\sigma z_{\frac{\alpha}{2}}}{\sqrt{n}}$ con confidenza $1-\alpha$. $\frac{\sigma z_{\frac{\alpha}{2}}}{\sqrt{n}}$ è la precisione.

Cosa rappresenta la confidenza, quantitativamente.

Esercizio 3 del 26/6/02. Esercizio 3 del 22/7/03: parte i), e introduzione alla ii). Per casa: cominciare a svolgere il compito del 23/9/03.

Lezione 23 (28/4, 2 ore ridotte ad una per assemblea di Facoltà). Dal teorema sugli intervalli di confidenza alla strategia per i test d'ipotesi (verifica della media di una gaussiana, con varianza nota). Significatività, modalità di esecuzione del test. Un esempio.

Lezione 24 (5/5, 2 ore). Esercizio 3.i del 4/6/02. Preannuncio delle diramazioni della teoria: test bilaterale/unilaterale, σ nota/incognita. Nozioni base di teoria dei test: ipotesi nulla H_0, ipotesi alternativa H_1, significatività α, regione critica C (sua dipendenza da α e dalla strategia del test), errore di I specie e sua probabilità ($=\alpha$), errore di seconda specie (solo definizione, per ora). Se un test è significativo per un certo α allora lo è per ogni $\alpha' > \alpha$. Si cerca allora il valore critico α_c tale che il test è significativo per ogni $\alpha > \alpha_c$ e non significativo per ogni $\alpha < \alpha_c$; lo chiamiamo anche "valore p". E' identificato dalla formula $|z| = q_{1-\frac{p}{2}}$, da cui $p = 2 - 2\Phi(|z|)$. Se si deve eseguire un test a livello α e si è calcolato il valore p, basta confrontare α con p.

Lezione 25 (8/5, 2 ore). Esercizio 5.iii del compitino. Esercizio: calc. $P_{\mu,\sigma^2}(X \in [\mu_0 - \delta, \mu_0 + \delta])$, dove si intende che $X \sim N(\mu, \sigma^2)$.

Errore di seconda specie: definizione, calcolo della sua probabilità. Potenza di un test. Esercizio 3.iii del 4/6/02. Osservazione: spesso un termine di β è quasi nullo.

Lezione 26 (12/5, 2 ore). Soluzione degli esercizi 4 e 5 del 13/6/03. Osservazioni: 4) si usa la varianza vecchia, ipotizzando che essa non sia cambiata (non esistono dati migliori sulla varianza); 5) una volta calcolato il valore p, per sapere l'esito del test con una data significatività α basta confrontare α con p.

Varianti. Prima variante: tast unilaterale invece che bilaterale. Dal teorema $P_{\mu_0,\sigma^2}\left(\sqrt{n}\frac{\bar{X}-\mu_0}{\sigma} > q_{1-\alpha}\right) = \alpha$ (dimostrato graficamente) si deduce la strategia unilaterale. $p = 1 - \Phi(z)$, $\beta(\mu) = \Phi\left(q_{1-\alpha} - \sqrt{n}\frac{\mu-\mu_0}{\sigma}\right)$. Per casa: es. 43 della lista.

Seconda variante: σ incognita. Si parte dalla teoria dell'intervallo di confidenza; introducendo le v.a. t di Student a n gradi di libertà ed i numeri ad esse associati $t_\alpha^{(n)}$ tabulati, vale il teorema $P_{\mu,\sigma^2}\left(\left|\sqrt{n}\frac{\bar{X}-\mu}{S}\right| \leq t_{\frac{\alpha}{2}}^{(n)}\right) = 1-\alpha$, da cui $\mu = \bar{X} \pm \frac{S \cdot t_{\frac{\alpha}{2}}^{(n)}}{\sqrt{n}}$ con probabilità $1-\alpha$. Per casa: es. 42 della lista (si deve usare il test con $t_{\frac{\alpha}{2}}^{(n)}$).

Lezione 27 (15/5, 2 ore). Esercizi vari. Un esercizio sull'approssimazione gaussiana. Un esercizio, sempre sull'approssimazione gaussiana, all'interno di una domanda circa un intervallo di confidenza per la media di una v.a. non gaussiana. Un esercizio sui test per la media, con varie diramazioni possibili (gaussiana e t di Student, bilaterale ed unilaterale).

Lezione 28 (19/5, 2 ore). Risoluzione esercizio sui test del compitino: valore p, potenza, numerosità necessaria per avere una certa potenza.

Il concetto di DOE (Design of Experiments, ovvero: progettazione di esperimenti), limitatamente al problema di progettare il *numero* di esperimenti che fornisce certi standards (es. una significatività ed una potenza prefissate, oppure una confidenza ed una precisione prefissate). La funzione potenza; β come funzione di $\left|\sqrt{n}\frac{\mu_0-\mu}{\sigma}\right|$; curve OC, loro uso grafico per il DOE.

Carte di controllo della qualità: partendo da μ_0 e σ noti inizialmente (e σ supposto sempre valido), prefissato $q_{1-\frac{\alpha}{2}}$ ad es. =3, scelta una numerosità n, si traccia la carta e si eseguono esperimenti a tempi regolari, eseguendo un test visivo ad ogni tempo.

Legge dei grandi numeri, con preliminari sulle disuguaglianze di Chebishev e Markov.

Lezione 29 (22/5, 2 ore). Introduzione alle idee della regressione lineare e non lineare: influenza di fattori su variabili, legame tra variabili, esemplificata attraverso l'esercizio n. 46 della lista. Calcolo della covarianza e dell'indice di correlazione sperimentali (cenno al legame tra indice positivo e retta approssimante con coefficiente angolare positivo, e viceversa per il negativo). Ricerca di un legame lineare del tipo $Y = \alpha X + \beta + e$. Esercizio: supponendo e a media nulla ed indipendente da X, calcolare $E[Y]$, $E[XY]$, $Cov(X,Y)$, in funzione di α, β, e momenti della X. Si arriva alle formule $\alpha = Cov(X,Y)/Var(X)$, $\beta = E[Y] - \alpha E[X]$. Se ne ricavano le formule per gli stimatori di α, β a partire da un campione sperimentale (x_1, y_1), ... , (x_n, y_n). Facoltativamente, si può vedere il metodo dei minimi quadrati per arrivare alle stesse espressioni per gli stimatori. Esempio dell'esercizio n. 46.

Cenno alla regressione nonlineare: metodo di linearizzazione.

Test per il parametro di una Bernoulli basato sulla distribuzione Binomiale (par. 8.6). Test di Fisher-Irwin (par. 8.6.1).