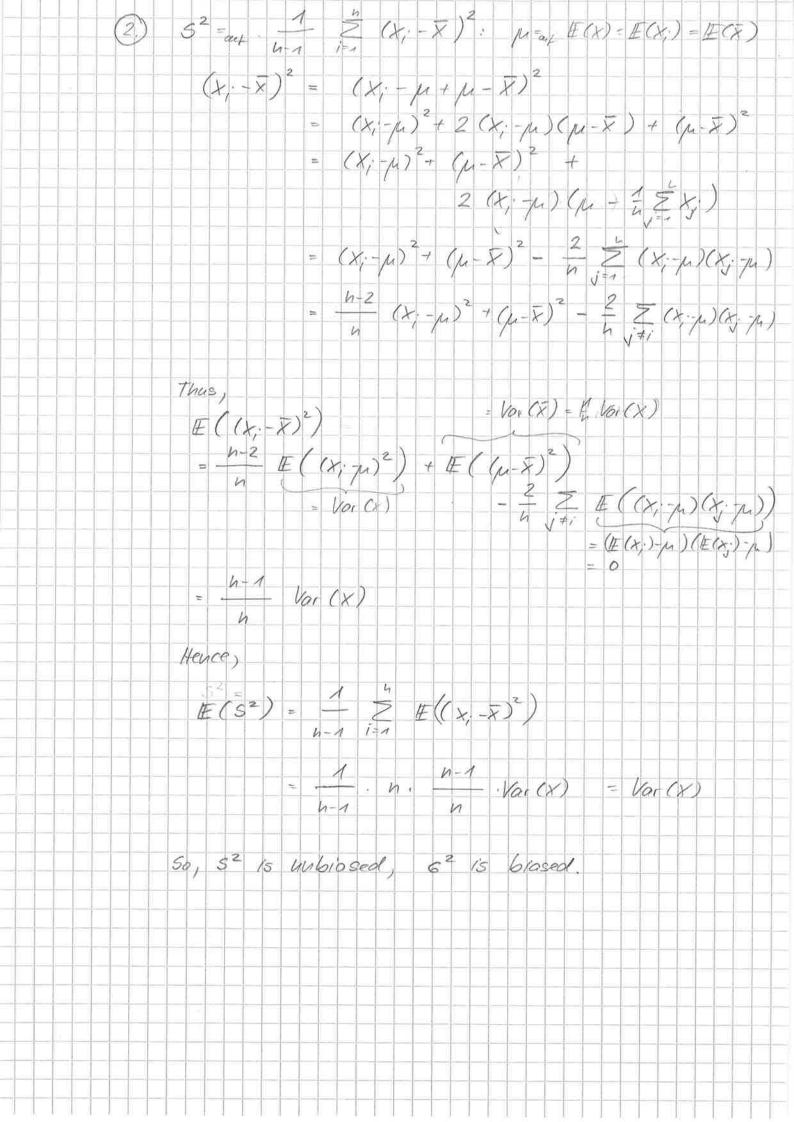
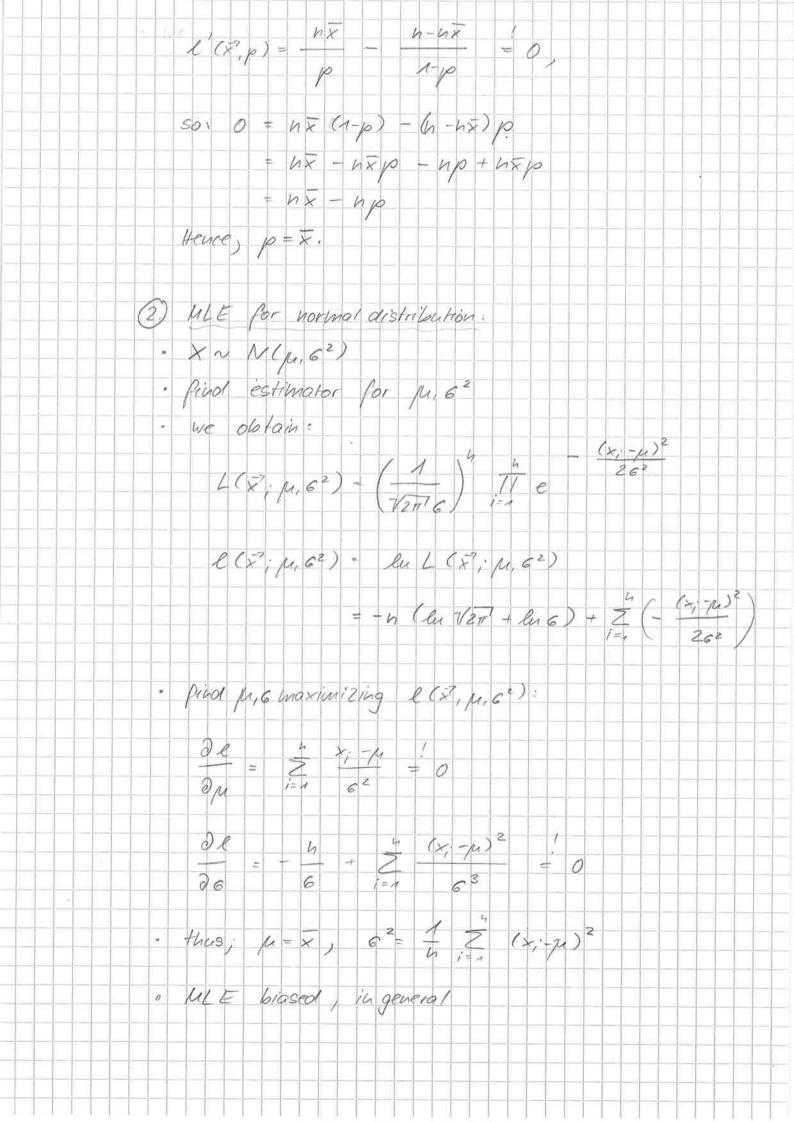
5. Inferential statistics - world based exploration of data, i.e., use of a prob model · goal: learn general statements from samples 5.1 Estimators Example: (Bernoulli distribution) · determine the failure perob, of hard-disk drive (charge) · X counts the humber of read/write access until first failure · P[x+k] = (1-p) p, i.e., x is geometricolly distrib. · it holds that I (x) = = 2 & (1-p) & 1p = p (- \(\frac{2}{k-1} (1-p)^{\frac{1}{k}} \) · Thus, p = E(x) · suppose, we have h drives, i.e., h r.v. x., ... , x. · consider X = at 1 2 x, · X is an estimator for # (X)

Definition 1. let X be a random variable with density fluction f(x, v). In estimator U for v of X is a random variable compased of Cindependent, identically distributed) sampling variables X1, Xn , i.e., U = la (+1, , x) Definition 2. Let U, U' be estimators of parameter v of X. (1) U is soid to be unbiosed if and only if I ((1) = v; otherwise, a is biosed. (2.) U is more efficient thou U if and only if $\mathbb{E}((u-v)^2) \leq \mathbb{E}((u'-v)^2).$ (3.) U is consistent if and only if # ((U-v)2)->0 pulses h -> 00; h is the sample side. Examples X = of to Z X; " We have $E(x) = E\left(\frac{1}{2} \frac{1}{2} x_i\right) = \frac{1}{2} \frac{1}{2} E(x_i) = \frac{1}{2} \frac{1}{2} E(x_i) + E(x_i)$ so, X is unbiased. $\mathbb{E}\left((X-V)^2\right) \stackrel{!}{=} Vo_1(X) = Vo_1\left(\frac{1}{n}\sum_{i=1}^n X_i\right)$ $= \frac{1}{n^2} \sum_{i=n}^{n} Vai(X_i) = \frac{1}{n} \cdot Var(X)$ So, X is consistent



How to construct istimators & Maximum-Litelihood estimators (MLE). X1, Xn is a sampling (independent, identically distr.) form a random vector & = (x1, ,x4) · X is distributed acc. to f(x; v) = Pol x = x J, i.e., X discrete · for X e R", define L(X, V) = at 1/1 f(x, iV) = 1/1 Pol X; = x;] - Pr L x = x 1 , X = x -- Lis called likelihood function · find v that maximiles ((x7, V), Definition 3. All estimator it for the parameter in f(x,v) of x is called haximum - Likelihood- Estimator (MLE) of a sampling & if and only it for all v, 1(x, v) \ \ (x, v) Example: @ (RILE for Berndilli distribution) = estimator for p st. P. [x; =1]=p, P. [x; =0]=1-p. for x? = (x1, ..., x1), Pp [x; =x;] = p x (1-p) 1-x; likelihood function: 4(xp)= 17 p * (1-p) 1-x; find p maximizing L, or find p maximizing land $\mathcal{L}(\vec{x}, p) = a_{i} \quad ln \quad \mathcal{L}(\vec{x}, p)$ $= \sum_{i=1}^{n} (x_{i} \cdot ln p + (1-x_{i}) ln (1-p))$ = hx lup + (n-nx) lu (1-p)



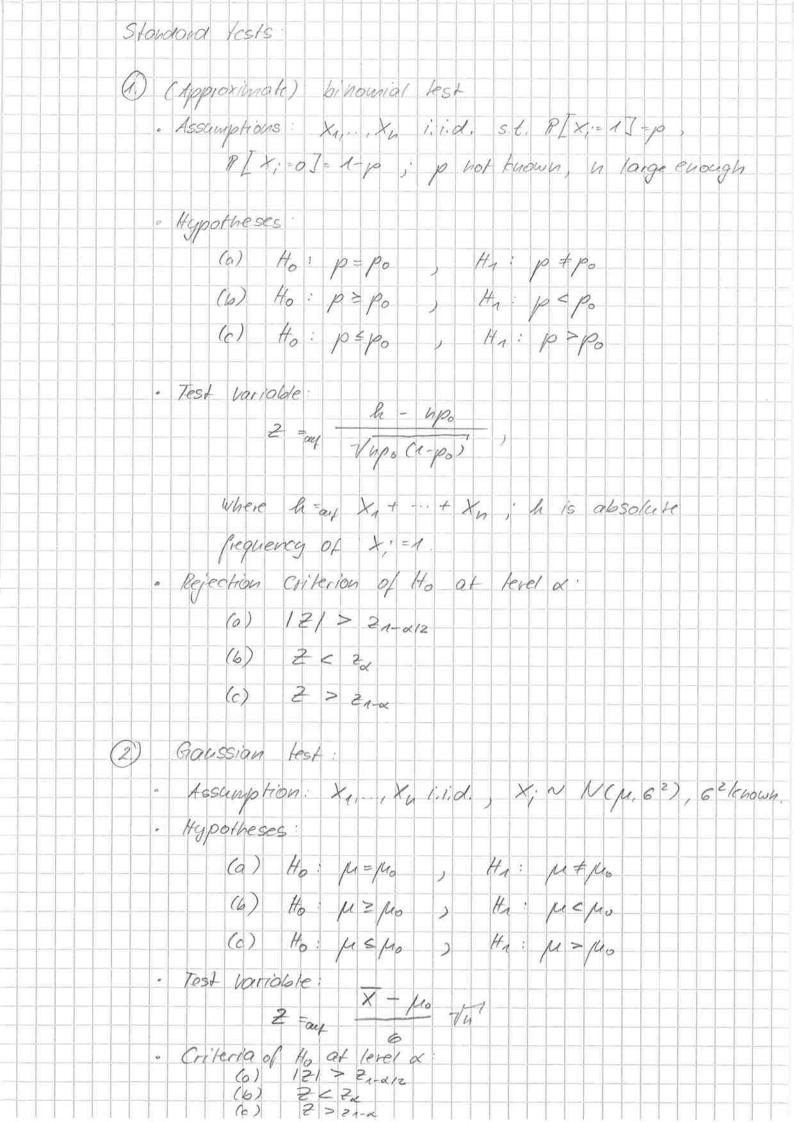
5.2 Confidence intervals · want to learn how good the estimation is use Itwo estimotors Un le St PLU, EVEU,]=1-0, · prob. 1- x is called confidence level . In, he I is called confidence interval; probabiling that & alle, as I is at most & · Appiral design: symmetrical confidence interval [4-5, 4+0] for d=0, i.e., C1=4-5, C2=4+5. Example: (normal distribution) · X N N(p, 62), X, , Xu sompling variables, · X N N (4, 62) · consider 2 = art 74 x - 14 ; so, 2 N N (0,1). · symmetrical confidence interval 1-c, c] for c>0 s.6. 1-a = PI-c = 2 = c7 $= P \left[\begin{array}{c} x + \frac{c6}{m} \leq \mu \leq x + \frac{c6}{m^2} \end{array} \right]$ · find c st. 1- x = P[-c = 2 = c]; PI-c==c]= \$(c)-\$(-c) - \$ (c) - (1 - \$ (c)) = 2 d (c) -1 = Thus, \$\overline{\pi} (c) = 1 - \frac{\pi}{2}, or, c = \overline{\pi} - 1 (1 - \frac{\pi}{2})

Definition 4. let x be a roudem variable with distribution 7 Then, xy eR S.t. Fx (xy)=y is called y-quantite of X. Example: For hormal alistribution, y-quantites are allenoted by zy. So, confidence interval for estimator for u of XNN(M,62): K = [X - 2(1-2) 6 , X + 2(1-2) 6 7

5.3 Hypothesis testing · want to learn if certain statements one true given a sample · e.g., is it we that pet for a Bernoulli disherbuted har. (1.) Definition of a lest in random yor. Xx,..., Xn, independent, identically distr. souple is a vector Z = (x1, +1) ER" critical set KSR for a hypothesis H K=aux & X / H is rejected given x & test variable T = le (X1, ..., X,) · R is decomposed into sets (intervals), associated with rejection or non-rejection; KER is associated with rejection · K = T - 1 (E) SRh Ho is called well hypothesis (which is to be examined) Hy is called alternative hypo (THA -> THO), often: Ha = THO e.g., Ho: p= = = = (or #0: p= = = An: p= =) with certain prob. and tests can fail, a is called significance level; typical values for x: 0,05, 0.01, 0.001 type - 1 error: Ho is true but is rejected Type-2 error. Ho is false but is failed to reject eg, for K=0: type-1 error = 0 (Ho hever rejected); type-2 error = 1 given a, construct T such that type-1 error is a

Example: Test for p of Bernoulli distribution Ho i $p = p_0$ $(H_0 = Lp_0, IJ)$, $H_1 : p < p_0$ $(H_1 = Lo, p_0)$ $(H_1 = Lo, p_0)$ $(H_2 = Lo, p_0)$ $(H_3 = Lo, p_0)$ $(H_4 = Lo, p_0)$ $(H_4 = Lo, p_0)$ · T is binomially distributed, TN B(n,p), i.e.,

PLT = & I = (h) p & (1-p) h-&; E(T) = hp, Var(T) = *pap) · de Moirre: lun - T - mp ~ N(0,1) · consider T = act Vnp (n-p) (approx. standard norm. distr.) · défermine significance level « Por q: type-1 error: max P. [Tek] = max P. [Teq] type-2 error max Pp IT & ET = max Pp, IT >97 we obtain: x = mox Pp, IT = Pp, IT = 9 - mpo I po I po (1-po) « Ø (9-npo Vopo(1po) ; O,1) thus, find a s.t. 9-hp = 20 So, 9 = 22 Vupo (1-po) + up



22 - lest: A: X1, X4 111d, Rx, = 21,..., 25 #: #o: P[x=i]=p, for all ie &1 ... , & 3 Ha: $P[X=i] \neq p$; for some $i \in \{1,...,k\}$ $T = a_{i+1} = a_{i+1}$ where his abs. frequency of X:=1' C: at level a: T > /2-1,1-x > the x2 - distribution; density purction is 1 (x) = 2 1 (k1) + e B