STATISTICS: Problem Set 3-Solutions

1. Consider an even number of independent Bernoulli-distributed observations where the two subsamples $Y_1, Y_2, .., Y_{n/2}$ and $Y_{n/2+1}, .., Y_n$ have parameters p_1 and p_2 respectively denoting $P(Y_i = 1)$.

(a) Work out the log-likelihood function for the whole sample, and show how you might construct the maximum likelihood estimators of p_1 and p_2.

(b) Using your result in (a) above, show that $Z_1 = \frac{2}{n}\sum_{i=1}^{n/2} Y_i \longrightarrow p_1$ and $Z_2 = \sum_{i=n/2+1}^{n} Y_i \longrightarrow p_2$, where $\longrightarrow$ denotes "convergence in probability".

(c) Combine the results in (b) to show that $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \longrightarrow \frac{p_1+p_2}{2}$, and find the maximum likelihood estimate of the parameter p for the whole sample(without taking into account the change in p).

(d) Comment briefly on why the Law of Large Numbers cannot be used for analyzing $\bar{Y}$ directly.

(e) Describe (but do not solve in detail) a likelihood ratio test of the null hypothesis $H_0 : p_1 = p_2$ against the alternative hypothesis $H_A : p_1 \neq p_2$.

2. The standardized bivariate normal distribution of two random variables X and Y with a correlation of ρ ($|\rho| < 1$) is characterized by the density:

$$f_{X,Y}(x, y; \rho) = (2\pi)^{-1}(1 - \rho^2)^{-\frac{1}{2}} exp\left[-\frac{x^2 - 2\rho xy + y^2}{2(1 - \rho^2)}\right],$$

for $-\infty < x, y < +\infty$.

(a) What are $E(XY)$ and $E(Y|X = x)$?

(b) Show that X and Y are independent if $\rho = 0$.

(c) Derive $log(f_{X,Y}(x, y; \rho))$ and differentiate this function with respect to ρ. Can you solve this result for ρ?

3. $[X,Y]' \sim f(x,y) = \frac{1}{\pi}\exp(-\frac{1}{2}(x^2+y^2))\mathbf{1}_A(x,y)$, $A = \{(x,y)|\ xy > 0\}$. Find marginal df of X and Y. What general conclusion can you draw from this exercise and from the fact, that marginal distributions of a bivariate Normal distribution are Normal?

4. A sample of n random variables $\{X_i\}$ is independently distributed as $N(\mu_X, \sigma_X^2)$.

a) Let $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ denote the sample average. Show that $E(\bar{X}) = \mu_X$, $E(\bar{X}-\mu_X)^2 = \frac{\sigma_X^2}{n}$ and hence that $\bar{X} \sim N(\mu_X, \sigma_X^2/n)$; discuss how this finding helps estimate μ_X when it is unknown.

b) Let $\nu^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$ be the sample variance of X around the sample mean. Show that $E(\nu^2) = \sigma_X^2$ and that $(n-1)\nu^2/\sigma_X^2$ is distributied as chi square with (n-1) degrees of freedom.

c) In general if variables $\{X_i\}$ are iid from a distribution with finite variance, and if $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$, then $cov(\bar{X}, Xi - \bar{X}) = 0$ for any r=1,2,...,n. Infer from this, that mean and sample variance of an iid sample from N(mu,sigma^2) are independent, and thus the statistic $t = \frac{\bar{X}-\mu_X}{(n-1)\nu^2}$ has Student's t distribution with (n-1) degrees of freedom.

d) Build an exact, small-sample confidence interval for μ_X