

Stat 471: Lecture 6

Multivariate normal distributions Part II.

Moo K. Chung
mchung@stat.wisc.edu

September 17, 2003

1. For given bivariate data $w_i = (x_i, y_i)$, we are interested in testing if the data follows bivariate normal. Bivariate normal data is somewhat simpler to analyze. If X and Y are bivariate normal with $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$,

$$E(Y|X = x) = \beta_0 + \beta_1 x,$$

where $\beta_1 = \rho\sigma_Y/\sigma_X$.

2. You may use χ^2 goodness-of-fit test or other tests to check bivariate normalness but a simpler approach would be to check if given bivariate data can be generated via linear transform

$$w = Hz + \mu \sim N(\mu, HH')$$

where $z \sim N(0, I)$. First estimate the sample covariance matrix:

$$\hat{V} = \frac{1}{n} \sum_{i=1}^n (w_i - \bar{w})(w_i - \bar{w})',$$

where $\bar{w}$ is the sample mean vector. The sample variance matrix is an estimator of unknown population covariance V . It can be computed using the *Kronecker tensor product* $\otimes$. For $A = (a_{ij})$ and B , $A \otimes B = (a_{ij}B)$.

```
>>barw=mean(w,2);
>>temp=w-kron(ones(1,1000),barw);
>>hatV=temp*temp'/1000;
hatV =
    2.0982    1.0232
    1.0232    1.9711
```

Let $\hat{V} = \hat{H}\hat{H}'$ be the Cholesky factors. Then $z = \hat{H}^{-1}w$.

```
>> hatH=chol(hatV) '
hatH =
    1.4485         0
    0.7064    1.2133
>> inv(hatH)
ans =
    0.6904         0
   -0.4019    0.8242
>> newz=inv(hatH)*w;
>>qqplot(newz(1,:));qqplot(newz(2,:));
```

3. Other factorization technique for covariance matrix V . Suppose V has positive eigenvalues λ_i and corresponding normalized eigenvectors e_i . Let $Q = (e_1, e_2, e_3)$. Then $QQ' = I$ and $V = Q \text{Diag}(\lambda_1, \lambda_2, \lambda_3)Q'$. So $V^{1/2} = Q \text{Diag}(\lambda_1^{1/2}, \lambda_2^{1/2}, \lambda_3^{1/2})Q'$. For $z \sim N(0, I)$, $w = V^{1/2}z \sim N(0, V)$.

```
>>V=[2 1
      1 2]
>> [Q D]=eig(V)
Q = -0.7071    0.7071
      0.7071    0.7071
D = 1 0
      0 3
>> Q*D.^(1/2)*Q'
ans = 1.3660    0.3660
      0.3660    1.3660
>> V^(1/2)
ans = 1.3660    0.3660
      0.3660    1.3660
>>w=V^(1/2)*[z1 z2] ';
```

It can be shown that generating bivariate normal this way gives the identical distribution as the Cholesky factorization (HW2).