170E Week 8 Discussion Notes

Colin Ni

December 2, 2025

Problem 1. (Midterm 2 Problem 3, rephrased) A bag contains 5 red, 3 yellow, and 2 white marbles. We uniformly randomly pick 4 balls without replacement. Let R and Y denote the number of red and yellow balls selected, respectively.

(a) Determine $\operatorname{supp}(R, Y)$.

(b) Find the joint density of (R, Y).

(c) Compute the marginal density f_R.

Solution. Let us start with (b). Observe that if we select r red and y yellow marbles, then we must select $4 - r - y$ white marbles. Thus

$$f_{(R,Y)}(r, y) = \frac{1}{\binom{10}{4}} \binom{5}{r} \binom{3}{y} \binom{2}{4 - r - y}.$$

For (a), we can use our answer for (b) as a guide. For $f_{(R,Y)}(r, y)$ to be positive, we need $r \in [5]$, $y \in [3]$ and $4 - r - y \in [2]$. Moreover, since we only select 4 balls, we in fact have $r \in [4]$. Thus

$$\operatorname{supp}(R, Y) = \{(r, y) \in [4] \times [3] \mid 4 - r - y \in [2]\}.$$

For (c), observe that if we select r red marbles, then we must select $4 - r$ yellow or white marbles. Thus

$$f_R(r) = \frac{1}{\binom{10}{4}} \binom{5}{r} \binom{5}{4 - r}. \qquad \square$$

Problem 2. (Midterm 2 Problem 4, rephrased) Let Y be a continuous random variable supported on $\mathbb{R}$. We define the *entropy* $\mathrm{H}(Y)$ of Y as

$$\mathrm{H}(Y) = -\int_{-\infty}^{\infty} f_Y(y) \log(f_Y(y)) \, \mathrm{d}y.$$

Fix $\mu, \sigma^2 \in \mathbb{R}$.

(a) Compute the entropy of $Y \sim \mathrm{Norm}(\mu, \sigma^2)$.

(b) Show that $Y \sim \mathrm{Norm}(\mu, \sigma^2)$ has maximal entropy among all random variables supported on $\mathbb{R}$ with mean μ and variance σ^2.

Solution. For (a), we directly compute

$$\begin{aligned}
\mathrm{H}(Y) &= -\int_{-\infty}^{\infty} f_Y(y)\log(f_Y(y))\,\mathrm{d}y \\
&= -\int_{-\infty}^{\infty} f_Y(y)\log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(y-\mu)^2\right)\right)\mathrm{d}y \\
&= -\int_{-\infty}^{\infty} f_Y(y)\left(-\frac{1}{2}\log(2\pi\sigma^2)-\frac{1}{2\sigma^2}(y-\mu)^2\right)\mathrm{d}y \\
&= \frac{1}{2}\log(2\pi\sigma^2)+\frac{1}{2\sigma^2}\mathbb{E}((Y-\mu)^2) \\
&= \frac{1}{2}\log(2\pi e\sigma^2),
\end{aligned}$$

where in the last equality we use that $\mu_Y = \mu$.

For (b), let X be a random variable supported on $\mathbb{R}$ with mean μ and variance σ^2. Recall from last week that the KL divergence $D_{\mathrm{KL}}(X \parallel Y)$ from X to Y is nonnegative. Thus

$$\begin{aligned}
0 &\le D_{\mathrm{KL}}(X \parallel Y) \\
&= \int_{-\infty}^{\infty} f_X(x)\log\left(\frac{f_X(x)}{f_Y(x)}\right)\mathrm{d}x \\
&= -\mathrm{H}(X) - \int_{-\infty}^{\infty} f_X(x)\log(f_Y(x))\,\mathrm{d}x \\
&= -\mathrm{H}(X) - \int_{-\infty}^{\infty} f_X(x)\log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)\right)\mathrm{d}x \\
&= -\mathrm{H}(X) + \frac{1}{2}\log(2\pi e\sigma^2) \\
&= -\mathrm{H}(X) + \mathrm{H}(Y),
\end{aligned}$$

where in the second-to-last equality we use that $\mu_X = \mu$ and $\sigma_X^2 = \sigma^2$ to reuse the computation from (a). Thus $\mathrm{H}(Y) \ge \mathrm{H}(X)$. $\square$

Problem 3. (Lecture notes Exercise 4.2, augmented) Let $\mu_X, \mu_Y \in \mathbb{R}$, $\sigma_X, \sigma_Y > 0$, and let $\rho_{X,Y} \in [-1,1]$. Our definition of the bivariate normal distribution is via the pdf

$$f_{(X,Y)}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{X,Y}^2}}\exp\left(-\frac{1}{2(1-\rho_{X,Y}^2)}Q(x,y)\right)$$

where

$$Q(x,y) = \left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 - 2\rho_{X,Y}\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}.$$

(a) Consider the change of variables

$$Z_X = \frac{X-\mu_X}{\sigma_X} \qquad \text{and} \qquad Z_Y = \frac{Y-\mu_Y}{\sigma_Y}.$$

Show that (Z_X, Z_Y) is bivariate normal with $\mu_{Z_X} = \mu_{Z_Y} = 0$ and $\sigma_{Z_X} = \sigma_{Z_Y} = 1$ and $\rho_{Z_X,Z_Y} = \rho_{X,Y}$.

(b) Consider the change of variables

$$U_X = \frac{Z_X - \rho_{X,Y} Z_Y}{\sqrt{1-\rho_{X,Y}^2}} \quad \text{and} \quad U_Y = Z_Y.$$

Show that (U_X, U_Y) is unit (means zero, variances one, zero correlation) bivariate normal.

(c) Show that for any bivariate normal, if $\rho_{X,Y} = 0$, then X and Y are independent.

(d) Deduce the means, variances, covariance, and correlation of X, Y.

Solution. For (a), an inverse transformation is

$$x(z_X, z_Y) = \sigma_X z_X + \mu_X \quad \text{and} \quad y(z_X, z_Y) = \sigma_Y z_Y + \mu_Y$$

which has Jacobian

$$J = \begin{pmatrix} \sigma_X & 0 \\ 0 & \sigma_Y \end{pmatrix}.$$

Thus by the bivariate change of variables theorem, we have

$$f_{(Z_X,Z_Y)}(z_X, z_Y) = |J| f_{(X,Y)}(x(z_X, z_Y), y(z_X, z_Y))$$
$$\frac{1}{2\pi\sqrt{1-\rho_{X,Y}^2}} \exp\left(-\frac{1}{2(1-\rho_{X,Y}^2)}(z_X^2 + z_Y^2 - 2\rho_{X,Y} z_X z_Y)\right).$$

This is the pdf of a bivariate normal with the claimed parameters.

For (b), an inverse transformation is

$$z_X(u_X, u_Y) = \sqrt{1-\rho_{X,Y}^2}\, u_X + \rho_{X,Y} u_Y \quad \text{and} \quad z_Y(u_X, u_Y) = u_Y$$

which has Jacobian

$$J = \begin{pmatrix} \sqrt{1-\rho_{X,Y}^2} & \rho_{X,Y} \\ 0 & 1 \end{pmatrix}.$$

Thus by the bivariate change of variables theorem, we have

$$\begin{aligned} f_{(U_X,U_Y)}(u_X, u_Y) &= |J| f_{(Z_X,Z_Y)}(z_X(u_X, u_Y), z_Y(u_X, u_Y)) \\ &= \frac{1}{2\pi} \exp\Bigg(- \frac{1}{2(1-\rho_{X,Y}^2)}((\sqrt{1-\rho_{X,Y}^2}\, u_X + \rho_{X,Y} u_Y)^2 + u_Y^2 \\ &\qquad - 2\rho_{X,Y}(\sqrt{1-\rho_{X,Y}^2}\, u_X + \rho_{X,Y} u_Y) u_Y)\Bigg) \\ &= \frac{1}{2\pi} \exp\left(-\frac{1}{2}(u_X^2 + u_Y^2)\right). \end{aligned}$$

This is the pdf of a unit bivariate normal.

For (c), we simply note that when $\rho_{X,Y} = 0$, the pdf is

$$f_{(X,Y)}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y} \exp\left(-\frac{1}{2}\left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right)\right).$$

On the other hand, the marginal pdf of X is

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f_{(X,Y)}(x, y)\,\mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left(-\frac{1}{2}\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right) \mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right)
\end{aligned}$$

and similarly for Y. Since the product of their marginals is their joint, X and Y are independent.

For (d), by (c), we have that U_X and U_Y are independent. We now backtrack to compute the desired quantities. For example,

$$\mathrm{Cov}(Z_X, Z_Y) = \mathrm{Cov}(\sqrt{1-\rho_{X,Y}^2}U_X + \rho_{X,Y}U_Y, U_Y) = \rho_{X,Y}$$

so

$$\mathrm{Cov}(X, Y) = \mathrm{Cov}(\sigma_X Z_X + \mu_X, \sigma_Y Z_Y + \mu_Y) = \sigma_X\sigma_Y\rho_{X,Y}.$$

(In other words, we express X, Y in terms of Z_X, Z_Y, and then in terms of U_X, U_Y, for which all the quantities are already known.) $\square$