Problem. Midterm and Final.

Midterm and final scores in a large class have an (approximately) bivariate normal dist'n w/ parameters:

	mean	sd
midterm	65	18
final	60	20

$\boxed{\text{correlation: } 0.76}$

Q: What is the "estimated" mean final score of the students who were above the mean on the midterm?

→: Let U be the midterm score, and let V be the final score

Our task is to find:

$$\mathbb{E}[V \mid U > \mu_U] = ?$$

Let X and Z be U and V in standard units, resp.

Our ancillary task is to find:

$$\mathbb{E}[Z \mid X > 0] = ?$$

$$\mathbb{E}[Z \mid X > 0] = \frac{\mathbb{E}[Z \cdot \mathbb{I}_{[X>0]}]}{\mathbb{P}[X>0]}$$

by def'n.

$$\boxed{\mathbb{E}[Z \mid X>0]} = \int_{\cancel{-\infty}\ 0}^{\infty} \boxed{\mathbb{E}[Z \mid X=x]} \cdot \boxed{f_X(x \mid X>0)\,dx}$$

Law of Total Probability

$\mathbb{E}[Z \mid X=x] = \rho x$ $\quad$ $Z \mid X=x \sim \text{Normal}(\rho x, 1-\rho^2)$

for $x>0$:

$$f_X(x \mid X>0)\,dx = \mathbb{P}[X \in (x, x+dx) \mid X>0]$$

$$= \frac{\mathbb{P}[X \in (x, x+dx) \text{ and } X>0]}{\mathbb{P}[X>0]}$$

$$= \frac{\mathbb{P}[X \in (x, x+dx)]}{\frac{1}{2}} = \frac{f_X(x)\,dx}{\frac{1}{2}}$$

$$= 2 f_X(x) = 2 \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$\mathbb{E}[Z \mid X>0] = \int_0^{\infty} \rho x \cdot 2 \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx$$

$$= \frac{2\rho}{\sqrt{2\pi}} \int_0^{\infty} x e^{-\frac{x^2}{2}}\,dx =$$

$u = \frac{x^2}{2}$

$du = \frac{1}{\cancel{2}} \cdot \cancel{2} \cdot x\,dx = x\,dx$

$$= \frac{2\rho}{\sqrt{2\pi}} \underbrace{\int_0^{\infty} e^{-u}\,du}_{=1} = \frac{2\rho}{\sqrt{2\pi}}.$$

In this problem:

$$\frac{1.52}{\sqrt{2\pi}} = \underline{0.6063923}$$

$\Rightarrow$ Our final answer: $60 + 20 \cdot 0.6063923 =$

$= 72.12785$

$\square$

Bivariate Normal in Matrix Notation.

Consider a bivariate normal pair (U, V).

In 2D: we place the means into a vector

$$\mu := \begin{pmatrix} \mu_U \\ \mu_V \end{pmatrix}$$

anything in $\mathbb{R}^2$

The variances/covariances are placed into a matrix:

$$\Sigma := \begin{bmatrix} \sigma_U^2 & \sigma_U \sigma_V \rho \\ \sigma_U \sigma_V \rho & \sigma_V^2 \end{bmatrix}$$

(positive definite)

In 1D:

$$f_U(u) = \frac{1}{\sigma_U \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{u-\mu_U}{\sigma_U}\right)^2}$$

Then, the joint density of (U,V) can be written as:

$$f_{U,V}(u,v) = \frac{1}{2\pi} \cdot \frac{1}{(\det(\Sigma))^{1/2}} \exp\left(-\frac{1}{2} \underbrace{\begin{pmatrix} u-\mu_U \\ v-\mu_V \end{pmatrix}^T}_{1\times 2} \underbrace{\Sigma^{-1}}_{2\times 2} \underbrace{\begin{pmatrix} u-\mu_U \\ v-\mu_V \end{pmatrix}}_{2\times 1} \right)$$

Multivariate Normal Density.

Let $\mathbb{X} = (X_1, X_2, \dots, X_p)^T$ be

$$\text{Normal}\left(\text{mean} = \mu = (\mu_1, \mu_2, \dots, \mu_p)^T,\ \Sigma = \begin{bmatrix} \sigma_1^2 & \text{Cov} & \text{Cov} & \text{Cov} \\ \text{Cov} & \ddots & & \vdots \\ \text{Cov} & & \ddots & \\ \text{Cov} & \cdots & & \sigma_n^2 \end{bmatrix}\right)$$

w/ Σ positive definite.

Then,

$$\boxed{f_{\mathbb{X}}(\underbrace{x_1, \dots, x_p}_{x}) = \frac{1}{(2\pi)^{p/2}} \cdot \frac{1}{(\det(\Sigma))^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right) \quad \text{for all } x \in \mathbb{R}^p}$$