CS 215 : Data Analysis and Interpretation

(Instructor : Suyash P. Awate)

Quiz (Closed Book)

Roll Number: ______________________________

Name: ______________________________

For all questions, if you feel that some information is missing, make justifiable assumptions, state them clearly, and answer the question.

Relevant Formulae

• Poisson: $P(k|\lambda) := \lambda^k \exp(-\lambda)/(k!)$

• Exponential: $P(x;\lambda) = \lambda \exp(-\lambda x); \forall x > 0$

• Univariate Gaussian:

$$P(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-0.5\frac{(x-\mu)^2}{\sigma^2}\right)$$

---

1. (25 points)

   Consider two random variables X_1 and X_2 with Gaussian probability density functions (PDFs) $G_1(x;\mu_1,\sigma_1^2)$ and $G_2(x;\mu_2,\sigma_2^2)$, where the set of parameter values is $\theta := \{\mu_1,\mu_2,\sigma_1,\sigma_2\}$.

   Define a PDF $P(x) \propto G_1(x;\mu_1,\sigma_1^2)G_2(x;\mu_2,\sigma_2^2)$, where the PDF $P(\cdot)$ is suitably normalized so that it integrates to 1.

   • [5 points] Mathematically prove that the PDF $P(\cdot)$ belongs to a family of distributions that we have studied in class.

   • [5 points] Mathematically derive the mean value associated with the PDF $P(\cdot)$, in terms of the parameters in θ.

   • [5 points] Mathematically derive the variance value associated with the PDF $P(\cdot)$, in terms of the parameters in θ.

   Product of two univariate Gaussians:

   $G(z;\mu_1,\sigma_1^2)G(z;\mu_2,\sigma_2^2) \propto G(z;\mu_3,\sigma_3^2)$ (via completion of squares in the exponent)

   where

   $$\mu_3 = \frac{\mu_1\sigma_2^2 + \mu_2\sigma_1^2}{\sigma_1^2+\sigma_2^2}; \sigma_3^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}$$

   Define a PDF $Q(x) \propto G_1(x;\mu_1,\sigma_1^2) + G_2(x;\mu_2,\sigma_2^2)$, where the PDF $Q(\cdot)$ is suitably normalized so that it integrates to 1.

   • [5 points] Mathematically derive the mean value associated with the PDF $Q(\cdot)$, in terms of the parameters in θ.

   • [5 points] Mathematically derive the variance value associated with the PDF $Q(\cdot)$, in terms of the parameters in θ.

   $Q(x) = 0.5G_1(x;\mu_1,\sigma_1^2) + 0.5G_2(x;\mu_2,\sigma_2^2)$

   Expectation of random variable associated with $Q(\cdot)$

   $= \int xQ(x)dx$

   $= 0.5(\mu_1+\mu_2)$ (using linearity of integration)

   Variance of $Q(x) \propto G_1(x;\mu_1,\sigma_1^2) + G_2(x;\mu_2,\sigma_2^2)$

   $= \int x^2 Q(x)dx - 0.25(\mu_1+\mu_2)^2$

   $= 0.5(\sigma_1^2+\mu_1^2) + 0.5(\sigma_2^2+\mu_2^2) - 0.25(\mu_1+\mu_2)^2$

   (using linearity of integration and Var$(X) = E[X^2] - (E[X])^2$)

   ---

2. (20 points)

   Consider two independent continuous random variables X and Y, both having a uniform PDF over $[0,1]$.

   • [5 points] Define random variable $Z := X + Y$. Mathematically derive the PDF of Z.

   This is a symmetric triangular PDF

https://statweb.stanford.edu/~susan/courses/s116/node114.html

---

• [5 points] Define random variable $W := X - Y$. Mathematically derive the PDF of W.

This is a symmetric triangular PDF

www.math.wm.edu/~leemis/chart/UDR/PDFs/StandarduniformStandardtriangular.pdf

---

• [5 points] Mathematically derive the covariance of Z and W.

Using bilinearity of covariance:

Cov$(X + Y, X - Y)$

$=$ Cov$(X, X) -$ Cov$(X, Y) +$ Cov$(Y, X) -$ Cov(Y, Y)

$= 1 - 0 + 0 - 1 = 0$

rng (0); N=1e4; X = rand (N,1); Y = rand (N,1); Z=X+Y; W=X-Y;

hist(Z,100); grid on, pause, close

hist(W,100); grid on, pause, close

plot(Z,W,'ro'), axis equal, grid on, pause, close

• [5 points] Are Z and W independent ? If so, prove it. If not, prove so.

Z and W aren't independent, because:

$P_Z(0.5) > 0$ and $P_W(0.6) > 0$, but $P_{Z,W}(0.5, 0.6) = 0 \neq P_Z(0.5)P_W(0.6)$

Intuitively, the joint PDF of (X, Y) (imagine a 2D distribution over the unit square) gets rotated to form the joint PDF of (Z, W) (which is diamond-shaped), and then knowing the value of Z does give some information about the value of W (through $P(W|Z)$) and this information differs for different values of Z.
