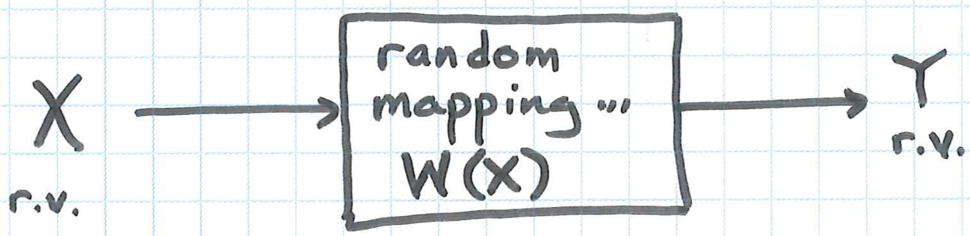
# ECE 645

J. V. Krogmeier

Session 2.B – January 22, 2021

FORM C
APPROVED FOR USE IN
PURDUE UNIVERSITY

# Joint Distributions of R.V.s




Observation $Y=y$

Goal: Learn something about $X$.

... $F_{X|Y}$, $f_{X|Y}$ ...

$P(X \in A \mid Y=y)$

...

mean, var cond. on $Y=y$

$\Longrightarrow$

$$F_{X,Y}(x,y) = P(X \le x, Y \le y)$$

Correlation, Covariance, joint pdf

$$F_{X|Y=y}(x|y) = P(X \le x \mid Y=y)$$

$$= \int_{-\infty}^{x} \underbrace{f_{X|Y}(\alpha|y)}_{f_{X,Y}(\alpha,y)/f_Y(y)} \, d\alpha$$

# Important Props of Conditional Expectation:

$$E\{X \mid Y=y\} \triangleq Q(y) \triangleq \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx$$

$\Longrightarrow$ Can create a new r.v. by $Q \triangleq Q(Y)$

$X, Y$ jointly distributed ~~r.v.~~ & $g(\cdot)$ a real-valued function s.t.

$$E\{|g(X)|\} < \infty$$

Then:

ⓐ $X \perp\!\!\!\perp Y \Longrightarrow E\{g(X)|Y\} = E\{g(X)\}$

ⓑ $X$ is a function of $Y$ ie $X = h(Y)$

$$E\{g(X)|Y\} = E\{g(h(Y))|Y\} = g(h(Y))$$

ⓒ $$E\{g(X)\} = E\{\underbrace{E\{g(X)|Y\}}\}$$

ⓓ $$E\{g(X)f(Y)|Y\} = E\{g(X)|Y\}f(Y)$$

FORM C
APPROVED FOR USE IN
PURDUE UNIVERSITY

# Proof © "Law of Total Expectation"

$$E\{g(X)\mid Y=y\} = \int_{-\infty}^{\infty} g(x)\, f_{X|Y}(x|y)\, dx$$ → this is a funct. of y. say $h(y)$.

If we subst. the r.v. Y into $h(y)$ get a new rv $h(Y)$.

$$E\{h(Y)\} = \int_{-\infty}^{\infty} h(y) f_Y(y)\, dy = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} g(x) f_{X|Y}(x|y)\,dx\right] f_Y(y)\, dy$$

Then use $f_{X,Y}(x,y) = f_{X|Y}(x|y)\, f_Y(y)$

$$\Rightarrow E\{h(Y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x) f_{X,Y}(x,y)\, dx\, dy = \int_{-\infty}^{\infty} g(x) \underbrace{\int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy}_{f_X(x)}\, dx$$

$$= \int_{-\infty}^{\infty} g(x) f_X(x)\, dx = E\{g(X)\}.$$

FORM C
APPROVED FOR USE IN
PURDUE UNIVERSITY

# Best Estimator Property

$X, Y$ jointly dist. rvs and $g(\cdot)$ a measurable function (i.e., so $g(X)$ is also a r.v.)

Assume all needed expectations exist. Consider problem of estimating $g(X)$ via functions of $Y$

$$\hat{\phi} = \underset{\phi \text{ Borel}}{\operatorname{argmin}} E\left\{ \left( g(X) - \phi(Y) \right)^2 \right\}$$

Then the MSE is minimized over all Borel functions $\phi$ by choosing

$$\hat{\phi}(Y) = E\{ g(X) \mid Y \}$$

FORM C
APPROVED FOR USE IN
PURDUE UNIVERSITY

# Proof of B.E.P.

$$E\{(g(X)-\phi(Y))^2\} = E\{\underbrace{E\{(g(X)-\phi(Y))^2 \mid Y\}}\}$$

$$= \int (g(x)-\phi(y))^2 f_{X|Y}(x|y)\,dx$$

$$= \int \left[ \quad \right] f_X(y)\,dy$$

**Lemma** $Z$ is a r.v. with finite mean

$$\hat{b} = \underset{b}{\operatorname{argmin}}\, E\{(Z-b)^2\}$$

$$\hat{b} = E\{Z\}.$$

The double integral is minimized if the interior integral is minimized pointwise for each $y$, ie if we pick $\phi(y)$ st.

$$E\{(g(X)-\phi(y))^2 \mid Y=y\} \text{ is minimized.}$$

Back to here : $\hat{\phi}(y) = E\{g(X) \mid Y=y\}.$

because $f_{X|Y}(x|y)$ is a valid pdf for $X$.

Table G.1 A Short Table of Q-Function Values

| $x$ | $Q(x)$ | $x$ | $Q(x)$ | $x$ | $Q(x)$ |
|---|---|---|---|---|---|
| 0 | 0.5 | 1.5 | 0.066807 | 3.0 | 0.0013499 |
| 0.1 | 0.46017 | 1.6 | 0.054799 | 3.1 | 0.00096760 |
| 0.2 | 0.42074 | 1.7 | 0.044565 | 3.2 | 0.00068714 |
| 0.3 | 0.38209 | 1.8 | 0.035930 | 3.3 | 0.00048342 |
| 0.4 | 0.34458 | 1.9 | 0.028717 | 3.4 | 0.00033693 |
| 0.5 | 0.30854 | 2.0 | 0.022750 | 3.5 | 0.00023263 |
| 0.6 | 0.27425 | 2.1 | 0.017864 | 3.6 | 0.00015911 |
| 0.7 | 0.24196 | 2.2 | 0.013903 | 3.7 | 0.00010780 |
| 0.8 | 0.21186 | 2.3 | 0.010724 | 3.8 | $7.2348 \times 10^{-5}$ |
| 0.9 | 0.18406 | 2.4 | 0.0081975 | 3.9 | $4.8096 \times 10^{-5}$ |
| 1.0 | 0.15866 | 2.5 | 0.0062097 | 4.0 | $3.1671 \times 10^{-5}$ |
| 1.1 | 0.13567 | 2.6 | 0.0046612 | 4.1 | $2.0658 \times 10^{-5}$ |
| 1.2 | 0.11507 | 2.7 | 0.0034670 | 4.2 | $1.3346 \times 10^{-5}$ |
| 1.3 | 0.096800 | 2.8 | 0.0025551 | 4.3 | $8.5399 \times 10^{-6}$ |
| 1.4 | 0.080757 | 2.9 | 0.0018658 | 4.4 | $5.4125 \times 10^{-6}$ |

Best estimator is a linear estimator in this case!

### 2.5.5 Conditional Expectation

+ Conditional expectation of $X$ given $Y = y$ is ...

$$\mathrm{E}[X|Y=y] = \begin{cases} \sum_x x p_{X|Y}(x|y) & \text{discrete} \\ \int x f_{X|Y}(x|y)\mathrm{d}x & \text{continuous} \end{cases}.$$

+ **Law of Total Expectation.** For rvs $X$, $Y$ ...

$$\mathrm{E}[X] = \mathrm{E}[\mathrm{E}[X|Y]].$$

## 2.6 Joint Gaussian Distribution

+ Two r.v.s $X$ and $Y$ are said to be bivariate normal (or Gaussian) if their joint pdf is

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \times \exp\left\{-\frac{1}{2(1-\rho^2)}F(x,y)\right\}$$

where $F(x,y)$ is the quadratic form:

$$\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2 - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}.$$

For such r.v.s

$$\begin{aligned} \mathrm{E}[X] &= \mu_x \\ \mathrm{E}[Y] &= \mu_y \\ \mathrm{Var}(X) &= \sigma_x^2 \\ \mathrm{Var}(Y) &= \sigma_y^2 \\ \mathrm{Cov}(X,Y) &= \rho\sigma_x\sigma_y \end{aligned}$$

where $-1 \le \rho \le +1$. (The bivariate normal distribution can be generalized to an arbitrary number of random variables. Such are called jointly Gaussian r.v.s.)

## 2.7 Gaussian Properties

+ Jointly Gaussian r.v.s $X$ and $Y$ are statistically independent if and only if (iff) they are uncorrelated, i.e., $\rho = 0$.

+ A linear combination of an arbitrary number of jointly Gaussian r.v.s is a Gaussian r.v.

+ **Conditional Gaussian:** Let r.v.s $X$ and $Y$ be jointly Gaussian. Then the conditional pdf of $X$ given $Y = y$ is a single variable Gaussian pdf with

$$\begin{aligned} \mathrm{E}(X|Y=y) &= \mu_x + \rho\frac{\sigma_x}{\sigma_y}(y-\mu_y) \quad * \\ \mathrm{Var}(X|Y=y) &= \sigma_x^2(1-\rho^2). \end{aligned}$$

+ **Gaussian Moments:** Let $X$ be a Gaussian random variable with mean zero and variance $\sigma^2$, i.e., $\mathcal{N}(0,\sigma^2)$. Then

$$\mathrm{E}(X^{2n}) = 1 \times 3 \times \cdots \times (2n-1)\sigma^{2n}$$

and

$$\mathrm{E}(X^{2n-1}) = 0$$

for $n = 1, 2, 3, \ldots$.

+ Let $X$ and $Y$ be zero mean jointly Gaussian r.v.s with equal variances $\sigma^2$ and $\rho = 0$ (i.e., they are statistically independent). Then the derived r.v.s $R = \sqrt{X^2+Y^2}$ and $\Theta = \arctan(Y/X)$ (four quadrant inverse tangent) are themselves statistically independent and $R$ is Rayleigh with parameter $\sigma$ and $\Theta$ is uniform on $[0, 2\pi)$.



6.