Econ241a: PS 4

Casey O'Hara

2017-10-21

Problems from handout

1.

Assume X and Y are jointly distributed with pdf $f(x,y) > 0$ for $(x,y) \in \mathcal{W}, \mathcal{W} \subset \mathbb{R}^2$. The marginals of X and Y are given by $f(x)$ with support $\mathcal{X}$ and $f(y)$ with support $\mathcal{Y}$. Define $g(X)$ as a function only of X. Prove that $E(g(x)) = \int_{x:x \in X} g(x)f(x)dx$.

$$\begin{aligned}
E(g(x)) &= \int_{x:x\in\mathcal{X}} \int_{y:y\in\mathcal{Y}} g(x)f(x,y)dydx && \text{(def of joint pdf, 4.1.10)} \\
&= \int_{x:x\in\mathcal{X}} g(x) \left[\int_{y:y\in\mathcal{Y}} f(x,y)dy\right] dx && \text{($g(x)$ independent of y)} \\
&= \int_{x:x\in\mathcal{X}} g(x)f_X(x)dx && \text{(marginal pdf $f_X(x) = \int_{y:y\in\mathcal{Y}} f(x,y)dy$, 4.1.3)} \blacksquare
\end{aligned}$$

2.

For the joint pmf in the table below:

	$x=1$	$x=2$	$x=3$
$y=0$	0.10	0.10	0.10
$y=1$	0.10	0.40	0.20

(a) Find the conditional expectation function $E(Y|X)$

This seems to be asking for $E(Y|X=x)$ for all x in the joint pmf.

$$E(g(Y)|X=x) = \sum_y g(y)f(y|x) \qquad \text{from def 4.2.3}$$

Rearrange to:

$$E(Y|X=x) = \sum_{y\in 0,1} y \times P_{Y|X=x}(y)$$

Calculate out for each x:

$$E(Y|X=x) = \begin{cases} (0\cdot 0.1 + 1\cdot 0.1)/0.2 = 0.50 & \text{for } x=1 \\ (0\cdot 0.1 + 1\cdot 0.4)/0.5 = 0.80 & \text{for } x=2 \\ (0\cdot 0.1 + 1\cdot 0.2)/0.3 = 0.33 & \text{for } x=3 \end{cases}$$

(b) Find the best linear predictor $E^*(Y|X)$

$h(x) = \alpha + \beta X$

$\hat{\beta} = Cov(X,Y)/Var(X)$

$Cov(X,Y) = E[XY] - \mu_X \mu_Y$

$E[X] = .2 * 1 + .5 * 2 + .3 * 3 = 2.1$

$E[Y] = 0 + 1 * (.1 + .4 + .2) = .7$

$E[XY] = (0)(1)(0.1) + (1)(1)(0.1) + (0)(2)(0.1) + (1)(2)(0.4) + (0)(3)(0.1) + (1)(3)(0.2)$

$E[XY] = .1 + .8 + .6 = 1.5$

$Cov(X,Y) = 1.5 - 2.1 * .7 = 0.03$

$Var(X) = E[X^2] - E[X]^2 = 1^2 * .2 + 2^2 * .5 + 3^2 * .3 - 2.1^2 = 0.49$

$\hat{\beta} = 0.03/0.49 = .0612$

$\hat{\alpha} = E[Y] - \hat{\beta}E[X]$

$= .7 - .0612 * 2.1 = .5715$

$E^*(Y|X) = h(x) = .5715 + .0612X$

(c) Prepare a table that gives $E(Y|x)$ and $E^*(Y|x)$ for $x = 1, 2, 3$.

	$x = 1$	$x = 2$	$x = 3$
$E(Y\|X)$	0.50	0.80	0.33
$E^*(Y\|X)$	0.6327	0.6939	0.7551

3.

Assume X and Y are jointly distributed with pdf $f(x,y) = x + xy$, $0 \leq x \leq 1$ and $0 \leq y \leq 1$. Define the bivariate random vector (U,V) as $U = X$ and $V = \sqrt{Y}$.

(a) Are X and Y independent?

Yes. Applying Lemma 4.2.7 to pdf $f(x,y) = x + xy$, we see

$$f(x,y) = x + xy = x(1+y) = g(x)h(y)$$

where $g(x) = x$ and $h(y) = 1 + y$.

(b) Are U and V independent?

Find the joint pdf of U and V.

First, find the inverse functions and the Jacobian:

$$
\begin{aligned}
x &= h_X(u,v) = U && \text{(where } U \in [0,1]\text{)}\\
y &= h_Y(u,v) = V^2 && \text{(where } V \in [0,1]\text{)}\\
|J| &= \begin{vmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & 2V \end{vmatrix} = 2V
\end{aligned}
$$

Then $f_{U,V}(u,v) = f_{X,Y}(h_x(u,v), h_y(u,v))|J|$ which translates to $2v(u+uv^2)$

This can be written as $f_{U,V}(u,v) = 2v(u+uv^2) = (2u)(v+v^3) = g(u)h(v)$, so U and V are independent.

(c) Find the marginal pdf of V.

$$
\begin{aligned}
f_V(v) &= \int_U f_{U,V}(u,v)du \\
&= 2(v+v^3)\int_0^1 udu \\
&= (v+v^3)u^2|_0^1 \\
&= v + v^3
\end{aligned}
$$

Problems from Casella and Berger:

4.19 (a)

Let X_1, X_2 be independent $n(0,1)$ random variables. Find the pdf of $(X_1 - X_2)^2/2$.

(Hint: What is the distribution of the square of a standard normal rv (Ch 2)? Does this result surprise you given that $X1$ and $X2$ are iid?)

Define $Z = (X_1 - X_2)^2/2$. Then $\sqrt{Z} = \frac{X_1 - X_2}{\sqrt{2}} = \frac{X_1}{\sqrt{2}} - \frac{X_2}{\sqrt{2}}$.

From Thm 4.2.14, the sum of two normals $\sqrt{Z} \sim n(\sum \mu_i, \sum \sigma_i^2) = n(0, 2\left(\frac{1}{\sqrt{2}}\right)^2) = n(0,1)$.

So $Z \sim n(0,1)^2$. The square of a normal distribution is χ^2 distribution, with one degree of freedom.

4.20

X_1, X_2 independent $n(0, \sigma^2)$ random variables.

(a) Find the joint distribution of Y_1 and Y_2, where

$$
Y_1 = X_1^2 + X_2^2 \quad \text{and} \quad Y_2 = \frac{X_1}{\sqrt{Y_1}}
$$

First, find joint pdf f_{x_1,x_2}: since X_1 and X_2 are independent normally distributed random variables,

$$
\begin{aligned}
f(x_1, x_2) &= f(x_1) \cdot f(x_2) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x_1^2}{2\sigma^2}} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x_2^2}{2\sigma^2}} \\
&= \frac{1}{2\pi\sigma^2} e^{-\frac{x_1^2+x_1^2}{2\sigma^2}}
\end{aligned}
$$

Next, find inverse functions of Y_1 and Y_2, and determine Jacobian:

$$
\begin{aligned}
h_{X1}(Y_1, Y_2) &= Y_2\sqrt{Y_1} \\
h_{X2}(Y_1, Y_2) &= \pm\sqrt{Y_1 - X_1^2} = \pm\sqrt{Y_1 - Y_1Y_2^2} \\
J = \begin{vmatrix} \partial h_{X1}/\partial y_1 & \partial h_{X1}/\partial y_2 \\ \partial h_{X2}/\partial y_1 & \partial h_{X2}/\partial y_2 \end{vmatrix} &= \begin{vmatrix} \frac{y_2}{2\sqrt{y_1}} & \sqrt{y} \\ \pm\frac{\sqrt{1-y_2^2}}{2\sqrt{y_1}} & \mp\frac{y_1y_2}{\sqrt{y_1-y_1y_2^2}} \end{vmatrix} = \pm\frac{1}{2\sqrt{1-y_2^2}}
\end{aligned}
$$

Note the $\pm$ comes from the fact that in h_{X_2}, lose one-to-one from X_2 to Y. So if we split the support from $X_2 \geq 0$ and $X_2 < 0$ and calculate each side separately we can use the transformation method since each side is monotonic.

Finally:

$$
\begin{aligned}
f_{Y_1,Y_2}(y_1, y_2) &= f_{X_1,X_2}(h_{X_1}(y_1, y_2), h_{X_2}(y_1, y_2)) \cdot (|J|_{x_2<0} + |J|_{x_2\geq 0}) \\
&= \frac{1}{2\pi\sigma^2} e^{-\frac{(y_2\sqrt{y_1})^2 + \sqrt{y_1 - y_1y_2^2}^2}{2\sigma^2}} \cdot (|J| + |-J|) \\
&= \frac{1}{2\pi\sigma^2} e^{-\frac{y_1}{2\sigma^2}} \cdot \frac{1}{\sqrt{1-y_2^2}}
\end{aligned}
$$

(b) Show that Y_1 and Y_2 are independent, and interpret this result geometrically.

We can divide $f_{Y_1,Y_2}(y_1, y_2)$ into $g(y_1)h(y_2)$, where

$$
g(y_1) = \frac{1}{2\pi\sigma^2} e^{-\frac{y_1}{2\sigma^2}} \quad \text{and} \quad h(y_2) = \frac{1}{\sqrt{1-y_2^2}}
$$

Therefore, Y_1 and Y_2 are independent.

Geometrically: $Y_1 = X_1^2 + X_2^2$ which is just the square of the Euclidean distance from the origin. Reframing $Y_2 = \frac{X_1}{\sqrt{Y_1}} = \frac{X_1}{\sqrt{X_1^2+X_2^2}}$ which is essentially the cosine of some angle. So this transformation is a modified version of polar coordinates, where instead of r and θ, we have r^2 and $\cos\theta$.

4.22

Let (X, Y) be a bivariate random vector with joint pdf $f(x, y)$. Let $U = aX + b$ and $V = cY + d$, where a, b, c, and d are fixed constants with $a > 0$ and $c > 0$. Show that the joint pdf of (U, V) is

$$f_{U,V}(u,v) = \frac{1}{ac} f(\frac{u-b}{a}, \frac{v-d}{c})$$

Find the inverse functions and Jacobian:

$$\begin{aligned}
x = h_X(u,v) &= \frac{U-b}{a} \\
y = h_Y(u,v) &= \frac{V-d}{c} \\
|J| &= \begin{vmatrix} \partial h_X/\partial u & \partial h_X/\partial u \\ \partial h_Y/\partial v & \partial h_Y/\partial v \end{vmatrix} = \begin{vmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{c} \end{vmatrix} = (\frac{1}{ac} - 0)
\end{aligned}$$

So the joint pdf is $f_{U,V}(u,v) = f(h_X(u,v), h_Y(u,v)) \cdot |J|$:

$$\begin{aligned}
f_{U,V}(u,v) &= f(h_X(u,v), h_Y(u,v)) \cdot |J| \\
f_{U,V}(u,v) &= \frac{1}{ac} \cdot f(\frac{u-b}{a}, \frac{v-d}{c}) \quad \blacksquare
\end{aligned}$$

4.26

X and Y are independent random variables with $X \sim$ exponential (λ) and $Y \sim$ exponential (μ). It is impossible to obtain direct observations of X and Y. Instead, we observe the random variables Z and W, where

$$Z = min\{X, Y\} \quad \text{and} \quad W = \begin{cases} 1 & \text{if } Z = X \\ 0 & \text{if } Z = Y. \end{cases}$$

(a) Find the joint distribution of Z and W.

We should break this into two pieces, one in which $W = 0$ and one in which $W = 1$. Start with $W = 0$, in which case we know $Z = \min\{X, Y\} = Y$ and $Y \leq X$.

Question: This problem does not specifically state we are to find a joint pdf; a check of the solutions manual instead shows how to determine the joint cdf. ***How do we know which to use and when, if not stated explicitly?***

$$
\begin{aligned}
F(Z,W) &= (Z \le z, W = 0) = P(Y \le z, Y \le X) \\
&= \int_0^z \int_y^\infty f(x,y)\,dxdy && \text{(Joint CDF from notes)} \\
& dx|_y^\infty \text{ first because evaluating the } Y \le X \text{ first} \\
&= \int_0^z \int_y^\infty f(x)f(y)\,dxdy && \text{(X, Y independent)} \\
&= \int_0^z \int_y^\infty \lambda e^{-\lambda x} \mu e^{-\mu y}\,dxdy && \text{(sub in pdfs)} \\
&= \int_0^z \mu e^{-\mu y} \cdot \int_y^\infty \lambda e^{-\lambda x}\,dxdy && \text{(pull out non-integrating constants)} \\
&= \int_0^z \mu e^{-\mu y}(-e^{-\lambda x})|_y^\infty dy \\
&= \int_0^z \mu e^{-(\mu+\lambda)y} dy \\
&= -\frac{\mu}{\mu+\lambda} e^{-(\mu+\lambda)y}|_0^z \\
F(Z,W) &= \frac{\mu}{\mu+\lambda}(1 - e^{-(\mu+\lambda)z}) && \text{(for case } W = 0\text{)}
\end{aligned}
$$

NOTE: I used the exponential form $\lambda e^{-\lambda x}$ but the solutions manual solves using $\frac{1}{\lambda} e^{-\frac{1}{\lambda}x}$. To make sure the solutions are identical, I verified that the forms are the same when I substitute in $\mu' = 1/\mu$ and $\lambda' = 1/\lambda$, i.e.

$$
\frac{\mu}{\mu+\lambda} = \frac{\lambda'}{\mu'+\lambda'}
$$

Following the exact same steps above, with a slight modification for the $W = 1$ case:

$$
\begin{aligned}
F(Z,W) &= (Z \le z, W = 1) = P(X \le z, X \le Y) \\
&= \int_0^z \int_x^\infty f(x)f(y)\,dydx \\
& \ldots
\end{aligned}
$$

I get:

$$
F(Z,W) = \frac{\lambda}{\mu+\lambda}(1 - e^{-(\mu+\lambda)z}) \qquad \text{(for case } W = 1\text{)}
$$

(b) Prove that Z and W are independent. (Hint: show that $P(Z \le z | W = i) = P(Z \le z)$ for $i = 0$ or 1.)

$$
\begin{aligned}
P(Z \le z) &= P(Z \le z | W = 1) + P(Z \le z | W = 0) \\
&= \left(\frac{\lambda}{\mu+\lambda} + \frac{\mu}{\mu+\lambda}\right)(1 - e^{-(\mu+\lambda)z}) \\
&= (1 - e^{-(\mu+\lambda)z})
\end{aligned}
$$

If independent, $P(Z \le z, W = i) = P(Z \le z)P(W = i)$. Find $P(W = i)$ for $i \in 0, 1$. Start with $W = 0$.

$$
\begin{aligned}
P(W=0) &= P(Y \leq X) \\
&= \int_0^\infty \int_y^\infty f(x)f(y)dxdy \\
&= \int_0^\infty \mu e^{-\mu y} \int_y^\infty \lambda e^{-\lambda x} dxdy \\
&= \int_0^\infty \mu e^{-(\mu+\lambda)y} dy \\
&= \frac{\mu}{\mu+\lambda}(0-1) = \frac{\mu}{\mu+\lambda}
\end{aligned}
$$

So for $W = 0$:

$$P(Z \leq z, W=0) = P(Z \leq z)P(W=0) = \frac{\mu}{\mu+\lambda}(1 - e^{-(\mu+\lambda)z})$$

and with the same steps, I find

$$P(Z \leq z, W=1) = P(Z \leq z)P(W=1) = \frac{\lambda}{\mu+\lambda}(1 - e^{-(\mu+\lambda)z})$$

Therefore, Z and W are independent for all $W \in 0, 1$.

■

4.30

Suppose the distribution of Y, conditional on $X = x$, is $n(x, x^2)$ and that the marginal distribution of X is uniform $(0, 1)$.

(a) Find $E(Y)$, $Var(Y)$, and $Cov(X, Y)$.

$$
\begin{aligned}
E(Y) &= E(E(Y|X)) && \text{(law of iterated expectations)} \\
E(Y|X) &= x && (\text{given } Y|X \sim n(x, x^2)) \\
E(Y) &= E(x) = (b+a)/2 = 1/2 && (x \sim u(0,1))
\end{aligned}
$$

$$
\begin{aligned}
Var(Y) &= E(Var(Y|X)) + Var(E(Y|X)) && \text{(defined in class notes 5p10)} \\
&= E(x^2) + Var(x) && (\text{given } Y|X \sim n(x, x^2)) \\
&= \int_0^1 x^2/(1-0)dx + (1-0)^2/12 = \frac{5}{12} && (X \sim u(0,1))
\end{aligned}
$$

$$
\begin{aligned}
Cov(X,Y) &= E(XY) - E(X)E(Y) && \text{(defined in class notes 5p10)} \\
&= E(E(XY|X)) - \frac{1}{2} \cdot \frac{1}{2} \\
&= E(E(X|X)E(Y|X)) - \frac{1}{4} \\
&= E(x^2) - \frac{1}{4} = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}
\end{aligned}
$$

(b) Prove that Y/X and X are independent.

(Hint for part b: does the pdf of $Y|x$ change for different values of x?)

Since $Y/X = x$ is basically $n(x/x, (x/x)^2) = n(1, 1)$, which does not involve x, then Y/X and X are independent.

4.44

Prove the following generalization of Theorem 4.5.6: For any random vector $(X_1, ..., X_n)$,

$$var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} var(X_i) + 2 \sum_{1 \le i < j \le n} cov(X_i, X_j)$$

Theorem 4.5.6 states that for X and Y are any two random variables and a and b are any two constants, then

$$var(aX + bY) = a^2 var(X) + b^2 var(Y) + 2ab \cdot cov(X, Y)$$

To generalize this proof, replace $aX + bY$ with $a_1X_1 + a_2X_2 + ... + a_nX_n = \sum_{i=1}^{n}(a_iX_i)$.

Helpful identities:

- mean of $aX + bY = \mathbb{E}(aX + bY) = a\mathbb{E}X + b\mathbb{E}Y = a\mu_x + b\mu_y$
- $Var(X) = \mathbb{E}(X - \mu_X)^2$
- $Cov(X, Y) = \mathbb{E}(X - \mu_X)(Y - \mu_Y)$

To show the generalization of Thm 4.5.6, proof:

$$\begin{aligned}
\text{mean of } (\sum_{i=1}^{n}(a_iX_i)) &= \mathbb{E}(\sum_{i=1}^{n}(a_iX_i)) = \sum_{i=1}^{n}(a_i\mu_{X_i})) && \text{(from mean identity)} \\
Var(\sum_{i=1}^{n}(a_iX_i)) &= \mathbb{E}(\sum_{i=1}^{n}(a_i(X_i - \mu_{X_i})))^2 && \text{(def of variance)} \\
&= \mathbb{E}((a_1X_1 - a_1\mu_{x_1}) + ... + (a_nX_n - a_n\mu_{X_n}))^2 && \text{(expand sum)} \\
&= \mathbb{E}(\sum_{i=1}^{n} a_i^2(X_i - \mu_{X_i})^2 + 2 \sum_{1 \le i < j \le n} a_ia_j(X_i - \mu_{X_i})(X_j - \mu_{X_j})) && \text{(square and rearrange)}
\end{aligned}$$

Note: The first sum is the sum of all elements multiplied by themselves; the second is the sum of each element times each other element, a la FOIL. Back to math:

$$\begin{aligned}
Var(\sum_{i=1}^{n}(a_iX_i)) &= \sum_{i=1}^{n} a_i^2\mathbb{E}(X_i - \mu_{X_i})^2 + 2 \sum_{1 \le i < j \le n} a_ia_j\mathbb{E}((X_i - \mu_{X_i})(X_j - \mu_{X_j})) && \text{(distribute } \mathbb{E}) \\
&= \sum_{i=1}^{n} a_i^2 Var(X_i) + 2 \sum_{1 \le i < j \le n} a_ia_j Cov(X_i, X_j) && \text{(def of } Var \text{ and } Cov)) \\
Var(\sum_{i=1}^{n}(a_iX_i)) &= \sum_{i=1}^{n} Var(X_i) + 2 \sum_{1 \le i < j \le n} Cov(X_i, X_j) && \text{(in given, all } a_i = 1)) \blacksquare
\end{aligned}$$

4.47

Let X, Y be independent n(0, 1) variables, nd define a new random variable Z by

$$Z = \begin{cases} X \text{ if } XY > 0 \\ -X \text{ if } XY < 0 \end{cases}$$

(a) Show that Z has a normal distribution.

The intuition here is to show that $P(Z < z) = P(X < z)$, since X is normally distributed. Because of the XY logic, we break this into $Z < 0$ and $Z > 0$ (since continuous, $P(Z = z) = 0\dots$). For $Z < 0$:

$$\begin{aligned}
P(Z < z) &= P(X < z \wedge XY > 0) + P(-X < z \wedge XY < 0) && \text{(account for case X < 0 and case X > 0)} \\
&= P(X < z \wedge Y < 0) + P(X > -z \wedge Y < 0) && \text{(if } XY > 0 \text{ and } X < 0,\ Y < 0\text{, vice versa)} \\
&= P(X < z)P(Y < 0) + P(X > -z)P(Y < 0) && (X, Y \text{ independent}) \\
&= P(X < z)P(Y < 0) + P(X < z)P(Y > 0) && (X, Y \text{ symmetric around 0, flip } <,>) \\
&= P(X < z)(P(Y < 0) + P(Y > 0)) && \text{(distributive)} \\
P(Z < z) &= P(X < z) && (P(Y < 0) + P(Y > 0) = 1)\blacksquare
\end{aligned}$$

Similar for $Z > 0$. Since distribution of Z in both partitions matches distribution of X $n(0, 1)$, Z is normal.

(b) Show that the joint distribution of Z and Y is not bivariate normal (hint: show that Z and Y always share the same sign)

Use a truthiness table (like a truth table, but not quite):

X	Y	XY	Z	Z, Y same signs?
+	+	+	+	TRUE
+	-	-	-	TRUE
-	+	-	+	TRUE
-	-	+	-	TRUE

Found online (and slightly modified for this problem): Two random variables Z and Y are said to be jointly normal if they can be expressed in the form

$$Z = aU + bV; \qquad Y = cU + dV$$

where U and V are independent normal random variables. In this case, if Z and Y always share the same sign, then $a + b = c + d \Rightarrow Z = Y$, which is not true (since only the sign of Z relates to Y, while the magnitude of Z relates to X), and doesn't seem to be in the spirit of jointly normal variables.

4.50

If (X, Y) has a bivariate normal pdf

$$f(x, y) = \frac{1}{2\pi(1-\rho^2)^{1/2}} exp\left(\frac{-1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right)$$

Show that $Corr(X, Y) = \rho$ and $Corr(X^2, Y^2) = \rho^2$.

$$Cor(X,Y) = \frac{Cov(X,Y)}{\sigma_X^2 \sigma_Y^2}$$

...

4.58 (a), (b) and (c).

For any two random variables X and Y with finite variances, prove that

(a) $Cov(X,Y) = Cov(X, E(Y|X))$

Proof:

$$\begin{aligned}
Cov(X,Y) &= E((X-\mu_x)(Y-\mu_y)) && \text{(by def of covariance, def 4.5.1)}\\
&= E(XY - Y\mu_x \; X\mu_y + \mu_x\mu_y) && \text{(multiply through)}\\
&= E(XY) - \mu_x E(Y) - \mu_y E(X) + \mu_x\mu_y && \text{(expectation is linear)}\\
&= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y) && (\mu_x = E(X), \mu_y = E(Y))\\
&= E(XY) + E(X)E(Y) && \text{(simplify expression)}\\
&= E(XY) + E(X)E(E(Y|X)) && \text{(law of iterated expectations)}\\
&= E(X \cdot E(Y|X) + E(X)E(E(Y|X)) && \text{(conditioning thm - notes 5p8)}\\
Cov(X,Y) &= Cov(X, E(Y|X)) && \text{(thm 4.5.3)} \blacksquare
\end{aligned}$$

NOTE: for that last step, for future reference: $Cov(X,Y) = E(XY) - E(X)E(Y)$, substitute $Y = E(Y|X)$.

(b) X and $Y - E(Y|X)$ are uncorrelated

If uncorrelated (not necessarily independent), then $Cov(X, Y - E(Y|X)) = 0$. To show $Cov(X, Y - E(Y|X)) = 0$, proof:

$$\begin{aligned}
Cov(X, Y - E(Y|X)) &= E[X \cdot (Y - E(Y|X))] - E(X)E(Y - E(Y|X)) && \text{(def of } Cov(\cdot)\text{, thm 4.5.3)}\\
&= E[XY] - E(X \cdot E(Y|X))] - E(X)E(Y) + E(X)E(E(Y|X)) && \text{(linearity of E)}\\
&= E[XY] - E(X \cdot Y)] - E(X)E(Y) + E(X)E(Y) && \text{(conditioning thm)} \square\\
Cov(X,Y) &= 0 && \text{(simplify)} \blacksquare
\end{aligned}$$

NOTE: not sure about the conditioning theorem line (noted by hollow sphere) where I assert that $E(Y|X) = Y$.

(c) $Var(Y - E(Y|X)) = E(Var(Y|X))$

Proof:

$Var(Y - E(Y|X)) = Var(Y) + Var(E(Y|X)) - 2Cov(Y, E(Y|X))$ (thm 4.5.6 expand $var(\cdot)$)

note to self: $a^2Var(X) + b^2Var(Y) + 2abCov(X, Y)$:

careful of signs!

$= Var(Y) + Var(E(Y|X)) - 2E(Y \cdot E(Y|X)) + 2E(Y)E(E(Y|X))$ (expand $Cov(\cdot)$ thm 4.5.3)

$= Var(Y) + Var(E(Y|X)) - 2E(E[Y \cdot E(Y|X)|X]) + 2E(Y)E(E(Y|X))$ (law of iterated E)

$= Var(Y) + Var(E(Y|X)) - 2E(E[Y|X]E[Y|X]) + 2E(Y)E(E(Y|X))$ (conditional of x?) $\square$

note to self: got from Jacob - check it

$= Var(Y) + Var(E(Y|X)) - 2E(E[Y|X]E[Y|X]) + 2E(E(Y|X))E(E(Y|X))$ (law of iterated E)

$= Var(Y) + Var(E(Y|X)) - 2[E(E[Y|X]^2) - E(E(Y|X))^2]$ (combine terms)

$= Var(Y) + Var(E(Y|X)) - 2Var(E[Y|X])$ (def of variance)

$= Var(Y) - Var(E(Y|X))$ (simplify)

$= E[Var(Y|X)]$ (law of total variance) $\blacksquare$

NOTE: there is one step in there that I was unable to confirm in the text or notes, noted by a hollow square.